INVESTIGATION OF A SULPHUR MOBILISATION GENE IN *PLASMODIUM FALCIPARUM*

William Richard Burns

Division of Parasitology

National Institute for Medical Research

The Ridgeway

Mill Hill

London NW7 1AA

Submitted for the degree of Doctor of Philosophy in University College London

(University of London)

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT.

Plasmodium falciparum is the principal causative agent of malaria. Cells of the organism harbour a plastid organelle that constitutes a potential drug target. Metabolic functions associated with the plastid include the synthesis of isoprenoids, fatty acids and haem. Iron-sulphur [Fe-S] clusters may also be assembled in the organelle. Two systems (ISC and SUF) are implicated in biological [Fe-S] cluster assembly in malaria. The ISC system probably functions in the mitochondrion while the SUF system operates in the plastid. The objective of this project was to investigate a gene in *P. falciparum* with marked homology to SufS, a cysteine desulphurase (EC 2.8.1.7) of the SUF system in *Escherichia coli* and *Arabidopsis thaliana*. *P. falciparum* SufS (PfSufS) is encoded on chromosome seven of the nuclear genome but is predicted to be trafficked to the plastid organelle. A number of techniques have been applied to investigate the function of PfSufS in this thesis. Following a general introduction in chapter 1, subsequent chapters describe progress with this work. In chapter 2, account is given of investigations into the properties of PfSufS using a combination of bioinformatic and *in vitro* techniques. Progress towards defining the intracellular targeting of PfSufS by transfection and immunolocalisation is described in chapter 3. Development of a simple colorimetric cysteine desulphurase assay is reported in chapter 4. Finally, attempts have been made to produce recombinant PfSufS for enzymological studies (chapter 5).

TABLE OF CONTENTS.

TABLE OF FIGURES AND TABLES.

ABBREVIATIONS.

$3C^{pro}$	Human rhinovirus 3C protease
AAT	Aspartate aminotransferase
ABC	Adenosine triphosphate binding cassette
ANN	Artificial neural network
At	*Arabidopsis thaliana*
BCA	Bicinchoninic acid
BIP	2,2′-bipyridyl
BLAST	Basic local alignment of sequences
ß-ME	ß-mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
BT	Bis-tris
CAT	Chloramphenicol acetyltransferase
cDNA	Copy deoxyribonucleic acid
CHEF	Contour clamped homogenous electric field
DDT	Dichlorodiphenyltrichloroethane
DMPD	N,N-dimethyl-*p*-phenylenediamine
DNA	Deoxyribonucleic acid
dS/dN	Ratio of rates of synonymous to non-synonymous mutation
DsRED	*Discosoma* red
DTT	Dithiothreitol
Ec	*Escherichia coli*
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
EES	Exo-erythrocytic schizogony
EGTA	Ethylene glycol-bis(2-aminoethylether)-N,N,N′,N′-tetraacetic acid
EM	Electron microscope
ER	Endoplasmic reticulum
ES	Erythrocytic schizogony
EST	Expressed sequence tag
FPLC	Fast protein liquid chromatography
GFP	Green fluorescent protein
GRAVY	Grand average of hydropathy
GSH	Reduced glutathione
GSSG	Oxidised glutathione
GST	Glutathione *S*-transferase
GuHCl	Guanidine hydrochloride
hDHFR	Human dihydrofolate reductase
HMM	Hidden Markov Model
HOMO	Highest energy occupied molecular orbital
HRP	Horseradish peroxidase
IB	Inclusion body
II	Instability index
INRA	Institut Nationale pour la Recherche Agronomique
IPTG	Isopropylthio-ß-D-galactoside
ISC	Iron-sulphur cluster assembly
KLH	Keyhole limpet haemocyanin
KO	Knock out
LMP	Low melting point
LUMO	Lowest energy unoccupied molecular orbital

Ma	*Mega-annus* (million years ago)
MB^+	Methylene blue
MBP	Maltose binding protein
MDCK	Madin-Darby canine kidney
MudPIT	Multidimensional protein identification technology
NCBI	National Center for Biotechnology Information
NIF	Nitrogen fixation
NIH	National Institutes of Health
NIMR	National Institute for Medical Research
NN	Neural network
NusA	N-utilising factor A
ORF	Open reading frame
PATS	Predict apicoplast targeted sequence
PBS	Phosphate buffered saline
PDB	Protein Data Bank
PEG 3350	Polyethylene glycol 3350
Pf	*Plasmodium falciparum*
PFGE	Pulsed field gel electrophoresis
PLP	Pyridoxal 5′-phosphate
PMP	Pyridoxamine 5′-phosphate
PMSF	Phenylmethylsulphonyl fluoride
PrP	Prion protein
RBC	Red blood cell
RNA	Ribonucleic acid
RNAP	RNAse protection
RPMI	Roswell Park Memorial Institute
RT-PCR	Reverse transcriptase-PCR
SDS	Sodium dodecyl sulphate
SGPP	Structural Genomics of Parasitic Protozoa
SOE-PCR	Splicing by overlap extension-PCR
SOR	Singlet oxygen resistance
SUF	Sulphur mobilisation
SVC	Second virial coefficient
TAIR	The Arabidopsis Information Resource
TCEP	Tris(2-carboxyethyl)phosphine
TE	Tris EDTA
THP	Tris(hydroxypropyl)phosphine
UCSF	University of California, San Francisco
WEHI	Walter and Eliza Hall Institute

ACKNOWLEDGEMENTS.

For practical assistance and useful discussions, I thank Drs. B. Clough, K. Rangachari, S. Sato, S. Adinolfi, J. W. Saldanha, D. H. Williamson, A. Petrovic, R. S. Buxton and D. Fernandez-Reyes (NIMR). For Sanger sequencing, I thank Mr. M. Strath (NIMR). For help with obtaining references, I thank the NIMR Library. For provision of reagents, I thank Drs. S. Lobréaux (INRA, Montpellier, France) and F. W. Outten (NIH, Bethesda, MA, USA). For supervision, I thank Drs. R. J. M. Wilson and M. J. Blackman. This project was funded by the Medical Research Council.

CHAPTER 1.

GENERAL INTRODUCTION.

Apicomplexan parasites are important human and animal pathogens. One genus in the group, *Plasmodium* spp., is the aetiological agent of malaria. Cells of *Plasmodium* spp. and related organisms harbour a plastid organelle. My thesis focuses on a malarial sulphur mobilisation gene predicted to encode a protein involved in iron-sulphur [Fe-S] cluster assembly within the plastid organelle. In three distinct sections, this introductory chapter will describe (1) the fundamentals of malaria; (2) the plastid organelle and the functions attributed to it; and (3) our knowledge of the mechanisms of [Fe-S] cluster assembly, based on work in bacteria, yeast and vascular plants.

1.1. FUNDAMENTALS OF MALARIA.

1.1.1. Malaria as a humanitarian catastrophe.

There is an old Italian proverb, "*malaria vien' dalla pentola*" ("malaria depends on what comes out of the pot").[1] Malaria is, indeed, a disease of poverty. Restricted to the developing countries of the tropics, it is estimated to cause 515 million clinical episodes per year; most of the burden falls in sub-Saharan Africa (70 %) and South East Asia (25 %).[2] The alleviation of such human misery is inevitably an important motivation for the scientific study of malaria. However, perhaps less appreciated is the remarkable biology of *Plasmodium* spp., the causative agents of the disease, and the intricate series of interactions they make with human host and mosquito vector.

1.1.2. *Plasmodium* spp., the aetiological agents of malaria.

1.1.2.1. Discovery of Plasmodium *spp.*

Periodic fevers resembling malaria have been known since ancient times.[3] Scientific work to identify the aetiological agent began in earnest after Pasteur (1822-1895) showed that microorganisms could cause disease. In 1879, Klebs (1834-1913) and Tommasi-Crudeli (1834-1900) attributed malaria to a soil bacterium named "*Bacillus malariae.*"[4] However, the true agent is eukaryotic, *Plasmodium* spp. (Apicomplexa: Haemosporida), and was discovered in 1880 by Laveran (1845-1922) in the blood of a French Legionnaire serving in Algeria.[5]

1.1.2.2. Taxonomy of Plasmodium *spp.*

Plasmodium spp. are members of a large (> 5000 spp.) phylum of unicellular eukaryotic microorganisms known as the Apicomplexa. DNA sequence analysis reveals the phylum to be an ancient lineage, falling among the chromalveolates. In a recent classification, Cavalier-Smith[6] demotes Apicomplexa to a subphylum of Miozoa, closely allied to Protalveolata and Dinozoa. The defining feature of the Apicomplexa is recognisable only in certain life cycle stages and at the cell ultrastructural level. This is the apical complex, composed of structures called polar rings, micronemes, rhoptries and (dependent upon species) a conoid and dense granules.[7] These structures are implicated in cell invasion (section 1.1.2.5). All members of the phylum are parasitic; several are important human and veterinary pathogens. Human malaria is attributed to one of four species of *Plasmodium*. *P. falciparum* is the most common cause of morbidity and mortality. In the following review, I will concentrate on this species.

1.1.2.3. Evolutionary history of P. falciparum.

DNA sequence analysis has facilitated investigation of the evolutionary history of *P. falciparum*. For example, comparison of 18S small subunit rRNA sequences from 11 *Plasmodium* spp. suggested that *P. falciparum* is most closely related to *P. reichenowi*, a parasite of the chimpanzee, *Pan troglodytes* (Primates: Hominidae); divergence apparently occurred ~ 6-10 million years (y) ago (Ma) (coincident with the divergence of the hominid hosts).[8] Several hypotheses are in currency over the chronology of subsequent evolution of *P. falciparum*.[9] One hypothesis ("Malaria's Eve") insists on a recent emergence from a population bottleneck ~ 5000 y ago; alongside the invention of agriculture and the rise of anthropophilic mosquito vectors.[10] An alternative hypothesis

postulates greater antiquity, contemporaneous with hominid population expansion from Africa; specifically, that extant *P. falciparum* arose *ca.* 100-180 thousand y ago.[11]

1.1.2.4. Cell ultrastructure of P. falciparum.

Clues to understanding the biology of *P. falciparum* have been gained by observation of parasites under the electron microscope (EM). The following brief discussion of cell morphology will focus exclusively on the erythrocytic stages of the parasite, as observed by Bannister, *et al.*[12] With regard to size, single merozoites are *ca.* 1.6 × 1.0 μm. Salient ultrastructural features common to all erythrocytic stages include: a nucleus; plastid organelle and mitochondrion (one of each per cell); an apical complex; microtubules; and, possibly, a Golgi apparatus. Given that the cell passes through a complex series of differentiations (sections 1.1.2.5 and 1.1.3.3), prominence of the various components varies through the lifecycle. For example, the apical complex is particularly important in the merozoite, an invasive stage of the parasite. Conversely, trophozoites and schizonts contain a pigment vacuole for sequestration of haemozoin derived from host haemoglobin.

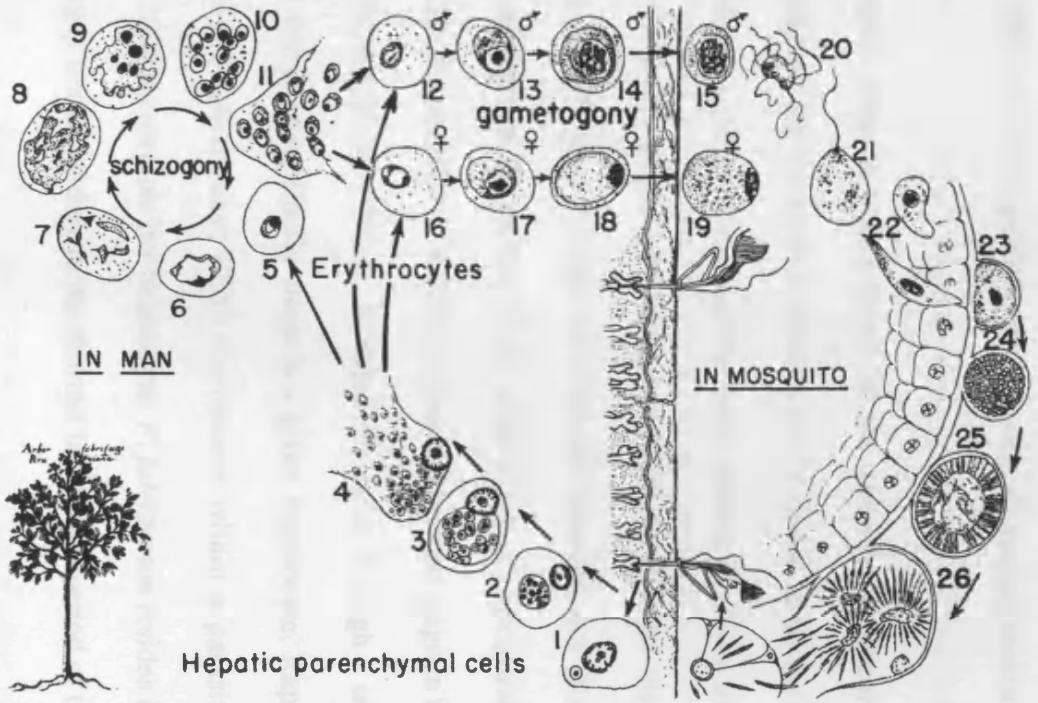
1.1.2.5. Malaria in the human host.

In humans, *P. falciparum* undergoes cell division in two sites, namely, the liver and the red blood cells (RBCs) (Fig. 1.1). In the following section, the events in the human host are briefly described, principally based on recent accounts given by Frevert;[15] and Bannister and Mitchell.[16]

FIGURE 1.1.

LIFE CYCLE OF *PLASMODIUM FALCIPARUM*.

P. falciparum undergoes a complex series of differentiations in the human and mosquito hosts. Following inoculation of the human by an infective mosquito bite, sporozoites invade liver cells and undergo exo-erythrocytic schizogony (1-4). Escape from the liver is followed by invasion of red blood cells, allowing erythrocytic schizogony (5-11); Transmission forms known as microgametocytes (12-15) and macrogametocytes (16-19) also develop in the blood. Ingestion by the mosquito triggers differentiation of these transmission stages: microgametocytes become microgametes (20) that fertilise the macrogamete (derived from the macrogametocyte) (21). Formation of motile stages known as ookinetes allows migration through the mosquito midgut wall (22). Encystment follows, giving oocysts (23). Rapid sporogony (24-25) leads to release of sporozoites (26) that migrate through the haemocoele and invade the salivary glands ready for inoculation. Drawing adapted from Fig. 4-1, p. 91 of Markell, *et al.*[13] © 1999, W.B. Saunders Company. ***Inset.*** Early illustration of *Cinchona* spp. (Rubiacae), the botanical source of the antimalarial drug (-)-quinine, from Jonston (1662) *Dendrographias*. Drawing from Plate 1 of Honigsbaum.[14] © 2001, Mark Honigsbaum.


9
10
8
11
12
13
14
15
20
gametogony
schizogony
21
7
16
17
18
19
22
23
6
5
Erythrocytes
IN MAN
IN MOSQUITO
24
25
4
3
2
26
1
Hepatic parenchymal cells

It is important to realise that little is know about the activity of *P. falciparum* as it moves towards, invades, and divides within, the human liver, owing to the difficulty of obtaining samples; instead, most data come from rodent and simian models. Conversely, *P. falciparum* erythrocytic stages are thoroughly studied, given facile maintenance *in vitro*.[17] Biological justification for this emphasis lies with the finding that 60 % of the *P. falciparum* genome is transcriptionally active in RBCs.[18] Cyclical infection of RBCs is also responsible for the clinical disease (section 1.1.2.6).

Infection begins when invasive forms of the parasite (sporozoites) are introduced subcutaneously as the mosquito vector takes a blood meal (< 25 sporozoites per probe[19]). Within *ca.* 30 min, sporozoites have entered the liver sinusoids. In patients with partial immunity, acquired by living in a region of endemic malaria, this process is hampered by immune attack. Invasion of hepatic macrophages (Kupffer cells) occurs, mediated by proteins on the surface of the parasite, e.g., *P. falciparum* circumsporozoite protein (PfCSP; PlasmoDB* PFC0210c). Sporozoites thus migrate through the Kupffer cells and enter the space of Disse. Restless migration through a series of hepatocytes ensues, until the parasite finally stops in a given hepatocyte. Rapid asexual division (exo-erythrocytic schizogony; EES) commences within a parasitophorous vacuole, producing ~ 1000 merozoites per hepatocyte. *P. falciparum* resides in the liver for *ca.* 6 d, approximately corresponding to the shortest incubation period of the disease.[20]

Mature merozoites escape the hepatocytes and rapidly invade erythrocytes; the invasion process is intricate but probably involves proteins released from the apical complex of the parasite. A second phase of cellular division (erythrocytic schizogony; ES) occurs, again within a parasitophorous vacuole. *Plasmodium* spp. express a large repertoire of

* The *Plasmodium* Genome Resource (http://plasmodb.org/)

variant surface antigens that are inserted into the RBC membrane; these can be recognised by the host immune system.

Parasites undergoing ES are classified morphologically as ring, trophozoite (feeding forms) or schizont (dividing form). Schizont rupture releases merozoites that invade fresh erythrocytes. In *P. falciparum,* ES is synchronised and typically lasts *ca.* 48 h, producing *ca.* 16 merozoites per schizont. Cycles of division and invasion proceed until the infection is: cleared by chemotherapy (section 1.1.2.7); controlled by the immune system; or the patient dies. A proportion (<< 10 %) of parasites differentiate into extracellular transmission stages (gametocytes) that are arrested in G_0 of the cell cycle and metabolically dormant. *P. falciparum* gametocytes are observed in peripheral blood within 7-15 d of the initial wave of asexual parasitaemia; circulation for long periods ($t_{1/2}$ = 2.4 d) prolongs the window of infectivity to mosquitoes.[21]

1.1.2.6. Clinical features of malaria.

Important clinical features of malaria have been reviewed by Miller, *et al.*[22] RBC lysis associated with schizont rupture typically manifests itself in paroxysms of fever. However, *P. falciparum* malaria can progress to a severe form, characterised by neurological and haematological sequelae. In endemic regions, severe malaria afflicts individuals lacking immunity, mainly children, and carries a case fatality rate of 10-20 %. A further complication may arise in pregnancy, whereby sequestration of parasitized RBCs in the placental circulation leads to foetal ischaemia. Vital to controlling these symptoms is the rapid knockdown of parasitaemia with antimalarial drugs (section 1.1.2.7).[23]

1.1.2.7. Control of malaria by antimalarial chemotherapy.

Synthetic analogues based on alkaloids of vascular plants comprise the bulk of the antimalarial pharmacopoeia.[24] Notably, the sesquiterpene lactones and the quinolines. Antimalarial use of sesquiterpene lactones can be traced to a Chinese herbal medicine (*qinghao*) recommended for the treatment of fever by the alchemist Ge Hong (AD 281-340).[25] Analysis of *Artemisia annua* (Asteraceae) disclosed the active principle to be artemisinin, a sesquiterpene 1,2,4-trioxane lactone. Parasiticidal action against *Plasmodium* spp. may require activation of the drug by Fe^{2+}; and interaction with specific cellular targets, e.g., a Ca^{2+}-dependent ATPase (PfATP6; PlasmoDB PFA0310c).[26] Quinolines, which may interfere with vesicle trafficking by binding to phospholipids,[27] were also identified from a botanical source. According to frescoes in the Santo Spirito Hospital (Rome, Italy), as early as 1638 the Countess of Chinchón, wife of the Viceroy of Lima (Peru), was cured of fever using an Indian febrifuge called *quinquina* ("bark of barks").[14] Bark of *Cinchona* spp. (Rubiaceae) indeed contains several alkaloids with antimalarial activity, most notably (-)-quinine, a quinoline-4-methanol.

1.1.3. *Anopheles* spp., the mosquito vectors of malaria.

1.1.3.1. Discovery of the mosquito vector.

Human malaria is exclusively transmitted by mosquitoes of the genus *Anopheles* (Diptera: Culicidae). Long suspected, notably by Manson (1844-1922) and Bignami (1862-1929), the role of the mosquito was conclusively demonstrated over a century ago.[28] In 1898, Manson's pupil, Ross (1857-1932), reported the transmission of bird malaria by culicine mosquitoes.[29] It fell to Bignami, Bastianelli (1862-1959) and Grassi

(1854-1925) to prove that anophelines were the vectors of human malaria.[30] It did not escape the notice of these pioneers that their discoveries immediately suggested a way to control the disease (section 1.1.3.4).

1.1.3.2. Natural history of Anopheles *spp.*

The anopheline vector system is complex and poorly understood. In sub-Saharan Africa, four groups probably make the largest contribution to disease transmission, namely, *A. funestus*, *A. gambiae*, *A. moucheti* and *A. nili*.[31] Mosquitoes need water in which to breed. Development of offspring is holometabolous and results in sexually mature adults within *ca.* 14 d. Both sexes feed on plant nectar, but only females are haematophagous (and thus, transmissive of malaria) as ovulation requires ingestion of blood.[32]

1.1.3.3. Malaria in the mosquito host.

The connection between malaria and the mosquito vector is extremely intimate and highly specific (Fig. 1.1).[33] Parasite development in the mosquito completes in *ca.* 14 d and begins with the ingestion of a small volume (1-2 μL) of blood containing a highly variable number of gametocytes (1-10^5). The mosquito midgut is the site of sexual recombination of malaria, which occurs after gametogenesis within 30-60 min of ingestion. Post fertilisation, motile ookinetes (5-6 per mosquito) migrate through the midgut epithelium and encyst under the basal lamina, forming oocysts (*ca.* 2 per mosquito). Notably at this stage, mosquitoes are able to mount an immune response against the parasite.[34] Extensive cell division in oocysts culminates in the release of

sporozoites (~ 8000 per oocyst), which migrate through the haemocoele to the salivary gland, ready for injection during a bloodmeal.

1.1.3.4. Control of malaria by destruction of the mosquito vector.

Malaria can be suppressed by killing the mosquito vector. Until the 1940s, attack was principally directed at larvae by physical destruction of breeding grounds and the use of larvicides, e.g., copper acetoarsenite (Paris Green). However, following the introduction of the organochlorine insecticide dichlorodiphenyltrichloroethane (DDT) by Müller (1899-1965), emphasis changed to the destruction of adult insects by house spraying.[35] Cheap personal protection measures are increasingly advocated, especially the avoidance of mosquito bites by sleeping under insecticide-treated bed nets.[36]

1.1.4. New avenues in the study and control of malaria.

Despite early successes, notably eradication in Europe,[37] malaria remains robustly endemic in many tropical countries. With so few classes of antimalarial available, rising drug resistance among parasites is of major concern.[38] Furthermore, insecticide resistance among mosquitoes has significantly reduced the efficacy of vector control.[39] Hence, new ways to combat malaria are continually being sought.

Planned interventions against malaria include novel chemotherapeutic agents;[40] a vaccine;[41] and large-scale release of transgenic mosquitoes refractory to transmission.[42] All these innovations have been included in a list of Grand Challenges in Global Health compiled by the Bill and Melinda Gates Foundation.[43]

While new avenues in research have been opened by sequencing the genomes of humans,[44] *P. falciparum*[45] and *A. gambiae*,[46] only limited progress has been made towards the development of a vaccine[47] and refractory mosquitoes.[48] Therefore, the most likely chance of short-term advancement lies with the development of new antimalarial drugs, notably, those targeting proteases, transporters, membrane biosynthesis, cell cycle control, the mitochondrion and the plastid organelle.[49] In order to develop drugs to exploit these targets, a greater understanding of basic parasite biology is required. The remainder of this chapter focuses on the plastid organelle (section 1.2) and a process newly associated with it, namely, iron-sulphur cluster assembly (section 1.3).

1.2. THE PLASTID ORGANELLE OF MALARIA.

1.2.1. Biology of the plastid organelle.

1.2.1.1. Discovery of the plastid organelle.

Four decades ago the plastid organelle of *Eimeria* spp. was observed by electron microscopy.[50] Subsequently, partitioning of *P. knowlesi* plastid DNA in a caesium chloride gradient was reported.[51] However, the significance of these observations was overlooked, and plastid features attributed to the mitochondrion, until the 1990s, when the existence of a distinct plastid organelle was deduced by Wilson (review of early data by Wilson, *et al.*[52]). Sequencing of the plastid genome of *P. falciparum* added support to this deduction,[53] while a connection was made to a specific intracellular compartment using *in situ* hybridisation techniques in *Toxoplasma gondii* (Apicomplexa: Coccidia)[54] and *P. falciparum.*[55] Plastids are found in many Apicomplexan taxa besides *Plasmodium* spp. and *T. gondii*, notably, *Theileria parva* (Piroplasmida)* and *Eimeria tenella* (Haemosporida).[56] However, plastids are absent from a distantly-related Apicomplexan, *Cryptosporidium parvum.*[57] Thus, besides the medical oddity of *Prototheca* spp. (Chlorellaceae),[58] Apicomplexan parasites are the only known group of plastid-bearing organisms to cause infectious human disease.

1.2.1.2. Ultrastructure of the plastid organelle.

Hopkins, *et al.*[59] used EM to examine plastid morphology in the erythrocytic stages of *P. falciparum* (Fig. 1.2).

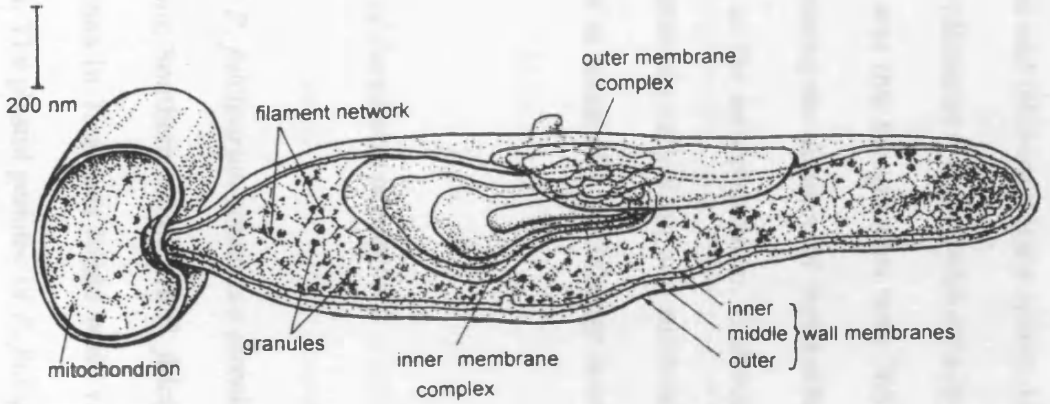
---

* Gardner, M.J., *et al.* (2005), *Science.*, in press.

FIGURE 1.2.

PLASTID ORGANELLE OF *PLASMODIUM FALCIPARUM*.

Drawing adapted from Fig. 9, p. 288 of Hopkins, *et al.*[59] © 1999, Urban & Fischer Verlag.


200 nm
outer membrane complex
filament network
inner
middle
outer
wall membranes
granules
inner membrane complex
mitochondrion

It was shown that each cell possessed a single plastid organelle of tubular shape. Organelle size was variable dependent on the developmental stage. For example, plastid dimensions in the merozoite were *ca.* 0.5 × 0.15 μm; equivalent measurements in the trophozoite were *ca.* 1.6 × 0.35 μm. Three bounding membranes were observed; although this number is controversial (discussed in chapter 3). Equally controversial was the observation of membrane invaginations, producing distinct complexes protruding into the lumen and inter-membrane spaces. Granular patches (diameter = 12 nm) linked by filaments (diameter = 4 nm) were seen in the plastid lumen. On the basis of size, one speculation was that the granules were 70S ribosomes while the filaments were DNA. It is worth noting the absence of thylakoids. In all erythrocytic stages, the plastid was juxtaposed to the mitochondrion. Contacts with the nuclear membrane, pigment vacuole, endoplasmic reticulum and microtubules were also noted. Taken together, these data hint at metabolic interchange between the plastid and other cell components.

1.2.1.3. Genome of the plastid organelle.

The plastid organelle of *P. falciparum* contains a circular genome of 35 kilobases (kb) Copy number is uncertain: Southern blotting in *P. falciparum* suggested one copy per plastid; while observations in *T. gondii* found wide variation (recently reviewed by Waller and McFadden[60]). The plastid genome of *P. falciparum* encodes: 25 tRNAs; 16S and 23S rRNA; three RNA polymerase genes; 17 ribosomal proteins; a translational elongation factor (PfEFTu; GenBank* CAA64593.1); a chaperone protein (PfClpC; GenBank CAA64596.1); and a member of the sulphur mobilisation (SUF) system (PfSufB; GenBank CAA64569.1).[53]

---

* GenBank (http://www.ncbi.nlm.nih.gov)

Additionally, there were seven short open reading frames (ORFs) of unknown function. Similar arrangements and genes were found on the plastid genomes of *T. gondii*[54] and *E. tenella*.[56]

1.2.2. Evolution of the plastid organelle.

1.2.2.1. Evolutionary history of the plastid organelle.

The Apicomplexan plastid shares a common ancestry with plant and algal plastids. However, delineating its true phylogenetic placement has proved difficult, due to the lack of photosynthetic marker genes, e.g., pigments. However, Apicomplexans likely share ancestry with dinoflagellates.[61,62] It is thus assumed that Apicomplexan plastids arose by secondary endosymbiosis.

In this context, it is useful to briefly summarise the proposed evolutionary history of the Apicomplexan plastid organelle. All plastid-bearing organisms probably originated by primary endosymbiosis, most likely of a cyanobacterium *ca.* 1500 Ma.[63] Divergence into three lineages followed: (1) red algae; (2) green algae and vascular plants; and (3) chromists.[64] Based on sequence analysis constrained by fossil evidence, divergence of red and green lineages is estimated to have occurred *ca.* 1340 Ma.[65] A second endosymbiotic event followed, leading to extant Miozoa (section 1.1.2.2).

In the Apicomplexa, it is uncertain whether secondary endosymbiosis was from the red or green lineage (reviewed by Funes, *et al.*[66]). Hints of red ancestry are found in the super-operon structure of the ribosomal proteins encoded on the plastid genome;[67] and

the sequences of plastid-targeted glyceraldeyhde-3-phosphate dehydrogenase (GAPDH) genes.[62] Recently it was shown that *T. gondii* amylopectin resembled the starch accumulated by red algae.[68] Conversely, the plastid EF-Tu sequence[54] and the nuclear location and structure of genes encoding subunit II of cytochrome *c* oxidase (COXII)[69] argue for a green lineage. However, it is worth noting that the validity of the COXII evidence has been challenged.[70] Recent analysis of the RNA polymerase genes encoded by the plastid genome of *E. tenella* also points to a green ancestry.[56]

1.2.2.2. Transposition of genes.

Cursory examination of the size of plastid genomes of extant eukaryotes, as compared to extant cyanobacteria, suggests that genes have been transposed. This is readily apparent for the plastid genome of *P. falciparum* (35 kb) as compared to the genome of, e.g., the cyanobacterium *Synechocystis* sp. PCC 6803 (3.57 Mb). Thus, concomitant with endosymbiosis, there must have been a significant transposition of DNA. Overwhelmingly, this would have been from endosymbiont genome to nuclear genome, given the minimal genome of extant plastids. Additionally, sequence analysis has been used to postulate a transfer of ribosomal protein genes from the mitochondrion to the plastid.[71] Lateral gene transfer from bacteria has also been postulated.[72]

Gauging the tempo of such genetic transpositions is difficult as rates probably vary. For example, the rates of chloroplast to nuclear gene transfer have been assessed in *Nicotiana tabacum* (Solanaceae) and *Chlamydomonas reinhardtii* (Chlamydomonadaceae) by tracking a reporter gene transfected into the chloroplast genome. In *N. tabacum*, transfer rates were high. Huang, *et al.*[73] detected one transposition event per 1.6×10^4 male gametes; using a different method, Stegemann, *et*

al.[74] detected one transposition event per 5×10^6 vegetative cells. Conversely, no transfer was detected in *C. reinhardtii.*[75]

1.2.3. Protein trafficking to the plastid.

Products of genes that now reside in the nucleus are imported back to the plastid where they act. Hence, the key to understanding the function of the plastid lies with identifying which gene products are trafficked to the plastid. Proteins are probably targeted to the plastid by an N-terminal bipartite leader sequence.[76] The length and amino acid composition of the plastid-targeting signal appears poorly conserved. Furthermore, one plastid-targeting signal can be shared by two different genes, as exemplified by stromal processing peptidase (PfSPP; GenBank AF453250) and porphobilinogen synthase (PfPBGS; PlasmoDB PF14_0381).[77] These factors make plastid-targeted proteins difficult to identify by conventional genome mining. Nevertheless, attempts have been made using *in silico* methods.[78,79] Such genome mining suggests that 545 nuclear-encoded proteins are imported to the plastid in *P. falciparum.*[80] Plastid trafficking is discussed in chapter 3.

1.2.4. Activity of the plastid organelle through the lifecycle.

1.2.4.1. Morphology of the plastid organelle through the lifecycle.

The morphology of the plastid organelle during ES has been extensively studied by light microscopy of *P. falciparum* transfected with plastid-targeted fluorescent markers, e.g., green fluorescent protein (GFP) (Sato and Wilson, NIMR, London, UK, unpublished observations). By these methods, it was shown that the plastid elongated

into a branched form during the trophozoite phase. This shape persisted through 2-3 nuclear divisions into the late schizont phase, when plastid segregation occurred. One plastid was faithfully bequeathed to each daughter cell.

In order to understand the basis of late segregation of the plastid, Striepen, *et al.*[81] studied plastid division in *T. gondii*, which benefits from unambiguous cell morphology. By this means, it was shown that the plastid organelle was connected to the nucleus, possibly via association with centrosomes. Consequently, it was suggested that the mitotic spindle and pellicle of the parent cell generated the forces necessary for division of the plastid. Plastid division factors, e.g., FtsZ and Min, have not been identified in the *P. falciparum* genome; hence, the molecular basis of division remains obscure. Also obscure is the morphology of the plastid in the EES and mosquito stages of the parasite. In this context, it is worth noting legacy EM observations of membranous structures in *P. falciparum* oocysts that are reminiscent of plastid organelles.[82]

1.2.4.2. Metabolic activity of the plastid organelle through the lifecycle.

Transcriptomic and proteomic studies offer a picture of plastid activity throughout the lifecycle of the parasite. Examination of these datasets provides a convenient, if rough, indication of plastid activity.* Erythrocytic stages of the parasite have been the most extensively studied. For example, Bozdech, *et al.*[18] showed that nuclear genes encoding plastid-associated proteins were expressed in all erythrocytic stages; while genes encoded on the plastid genome were upregulated in schizonts (33-36 h), coincident with

* Published transcriptomic and proteomic data were interrogated using the DGET function in Microsoft Excel. Interrogation was done using the PlasmoDB codes for the 545 putative plastid-targeted *P. falciparum* proteins generated by Ralph, *et al.*[80]

cell division. Le Roch, *et al.*[83] additionally demonstrated transcriptional activity of plastid-associated genes in the merozoite and sporozoite stages; activity was even observed in the metabolically dormant gametocyte stage. It is also worth noting that a microarray study of *P. falciparum* erythrocytic stages isolated from the venous blood of five malaria patients suggested that the plastid was active.[84]

Given the discrepancy between levels of mRNA and protein in *P. falciparum*,[85] it is vital to note additional evidence of plastid activity from proteomic studies. Analysis by multidimensional protein identification technology (MudPIT) of the erythrocytic stages of *P. falciparum* revealed a large number (234) of putative plastid-targeted proteins.[86] These data were confirmed by another proteomic study that also detected plastid-targeted proteins in gametocytes.[87]

Information on the activity of the plastid in EES and mosquito stages of the parasite is lacking. However, inferences can be drawn. For example, Wang, *et al.*[88] conducted an expressed sequence tag (EST) analysis of axenic pre-erythrocytic stages of *P. yoelii* and showed transcription of plastid-associated genes. Hints of plastid activity in gametes were provided by a mass spectroscopic study in *P. falciparum*.[87] Additional evidence for plastid activity in the ookinete, oocyst and salivary gland sporozoite stages was provided by MudPIT of *P. berghei* passaged through *A. stephensi* mosquitoes.[89] Taken together, these data suggest that the plastid is active throughout the malaria lifecycle.

1.2.5. Metabolic activities of the plastid organelle.

1.2.5.1. Functional analysis of the plastid by genome mining.

There is no genomic, transcriptomic, proteomic, structural or biochemical evidence for photosynthesis in the Apicomplexan plastid; hardly surprising, given the tenebrous lifecycle of *P. falciparum* and related organisms (Drs. B. Clough and R. J. M. Wilson, NIMR, pers. comm.). Therefore, intense efforts are being made to find other metabolic roles for the plastid organelle. One major block to progress has been the inability to fractionate Apicomplexan plastids, thus preventing biochemical and proteomic analysis. Hence, the only applicable method has been genome mining. Constrained by knowledge of heterotrophic plastids of vascular plants,[90] this entails the assembly of metabolic pathways *in silico* by reference to the plastid genome, and the set of 545 putative plastid-targeted proteins encoded by the nuclear genome. Such analysis has given some insights, but very few of the predictions have experimental support. I have structured this section principally according to the exhaustive *in silico* paper by Ralph, *et al.*[80]

1.2.5.2. Information processing.

One major activity of the plastid is the replication of the plastid genome.[91] Sequencing of the genes encoded by the plastid genome additionally suggested the plastid to be a minor site of transcription and translation.[53] Indeed, active transcription has been demonstrated.[18] Evidence for protein synthesis in the plastid has also been found.[92,93]

1.2.5.3. Energy, carbon and reducing power.

For anabolic biosynthesis to occur, energy, carbon and reducing power are required. Remarkably, mechanisms of energy generation remain obscure in the malaria plastid. Carbon acquisition pathways are modelled by reference to the heterotrophic plastids of vascular plants. In such plastids, Calvin cycle chemistry is not available and so carbon may be obtained by metabolism of precursors imported from the cytosol.[94] Appropriate transporters and enzymes for this process are apparently present in the nuclear genome of *P. falciparum* and are predicted to target the plastid organelle.[80] Reducing power could be supplied by ferredoxin (PfFDX; PlasmoDB MAL13P1.95), recycled by the action of ferredoxin $NADP^+$ reductase (PfFNR; PlasmoDB MAL6P1.197); both proteins have been shown to target the plastid in *T. gondii*.[95]

1.2.5.4. Fatty acid biosynthesis.

Fatty acid biosynthesis was indicated by acetate incorporation in *P. falciparum*.[96] Indeed, the complete pathway of type II fatty acid biosynthesis can be reconstituted *in silico*; enzymes of the pathway are predicted to target the plastid. These data have been strongly supported by targeting experiments. Notably, the plastid was targeted by acyl carrier protein (PfACP; PlasmoDB PFB0385w);[97] acetyl-CoA carboxylase (PfACCase);[98] and the pyruvate dehydrogenase complex (PfPDHC).[99] It is known that activation of PfPDHC requires lipoylation. Critically, a lipoic acid synthesis system was also localised to the plastid in *P. falciparum*[100] and *T. gondii*.[101] Evidence for the pathway also comes from the use of inhibitors. For example, it was noted that graminaceous weeds infesting broadleaf crops are selectively killed with aryloxyphenoxypropionic herbicides, via non-competitive inhibition of ACCase.[102] As

Apicomplexan ACCase shares features with susceptible plant enzymes, it was unsurprising that such herbicides were inhibitory to the growth of *P. falciparum* and *T. gondii*.[98,103]

1.2.5.5. Isoprenoid biosynthesis.

Isoprenoid synthesis in *P. falciparum* was recently demonstrated.[104] Genome mining indicates the pathway to be of type II (mevalonate independent), as found in bacteria and the chloroplasts of vascular plants.[105,106] Furthermore, it is predicted to be plastid-targeted.[80] Indeed, there is some evidence for plastid targeting, notably, the plastid localisation of 1-deoxy-D-xylulose 5-phosphate reductoisomerase (PfDXR; PlasmoDB PF14_0641) in *T. gondii*[107] and 4-diphosphocytidyl-2C-methyl-D-erythritol kinase (PfIspE; PlasmoDB PFE0150c) in *P. falciparum*.[108]

1.2.5.6. Haem biosynthesis.

Most of the enzymes of haem biosynthesis can be identified in the *P. falciparum* nuclear genome.[80] However, it appears that steps of the pathway are shared between mitochondrion, plastid and cytosol; an unprecedented division of labour. Critical enzymes of the pathway have been shown to target the plastid, notably PfPBGS,[109] hydroxymethylbilane synthase (PfHMBS; PlasmoDB PFL0480w)[109] and ferrochelatase (PfFC; PlasmoDB MAL13P1.326).[110] Additionally, complementation experiments in *E. coli* strongly suggest that PfPBGS[111] and PfFC[112] are enzymatically active.

1.2.6. Exploitation of the plastid organelle as a drug target.

1.2.6.1. The plastid organelle is essential and is susceptible to drug attack.

A working plastid is essential for the survival of Apicomplexan organisms and therefore constitutes a potential drug target.[113] Given the similarity between bacterial and plastid pathways, many existing antibiotics inhibit the plastid and are active against malaria.[114] For example, clindamycin, macrolides, chloramphenicol and doxycycline most likely attack the ribosomal subunits of the plastid organelle.[115] Doxycycline is indicated for antimalarial chemoprophylaxis.[116] The remainder of this section describes two cases where novel antimalarial drugs have been identified rationally from an understanding of plastid metabolism.

1.2.6.2. Attack on type II fatty acid biosynthesis.

Type II fatty acid biosynthesis constitutes an important potential target for antimalarial action as it is absent from the human host.[117] Greatest success has been achieved against enoyl ACP reductase (PfENR; PlasmoDB PFF0730c). This is due to the availability of an inhibitor, triclosan, patented as a biocide by Ciba-Geigy in 1966.[118] Triclosan was recently shown to be parasiticidal to *P. falciparum* erythrocytic stages *in vitro* (IC_{50} = 0.7 μM).[96,119] Triclosan is also parasiticidal to *T. gondii*;[120] confirmation of action in the *T. gondii* plastid was achieved by attachment of the drug to a rhodamine-conjugated octo-arginine oligomer, thus allowing localisation in live *T. gondii*.[121] In *P. berghei*-infected rodents, a single bolus of triclosan (38 $mg.kg^{-1}$) achieved radical cure of malaria.[96]

1.2.6.3. Attack on type II isoprenoid biosynthesis.

Absence from the human host makes type II isoprenoid biosynthesis another compelling target for drug design.[122] To my knowledge, this is the only target exploited in the clinical setting on the basis of rational knowledge of the plastid. Indeed, there has been marked success against PfDXR using fosmidomycin, a phosphonic acid antibiotic isolated from *Streptomyces lavendulae.*[123] It has antimalarial properties *in vitro* and in animal models.[107] Furthermore, clinical evaluation in Gabon[124,125] and Thailand[125] indicated the value of the drug for treating *P. falciparum* infection. The drug was of greatest clinical value in a 3:1 ratio with clindamycin.[126,127] Recall that clindamycin is also known to target the plastid.[115] These data validate the contention (section 1.1.4) that an understanding of basic plastid biology can rapidly lead to clinical innovation.

1.2.7. The plastid organelle as a site of iron-sulphur cluster assembly.

[Fe-S] clusters are ubiquitous and extremely versatile prosthetic groups.[128] The plastid in malaria is likely to have a high demand for [Fe-S] clusters. For example, ferredoxin[95] and lipoic acid synthase,[101] both [Fe-S] cluster-containing proteins, target the plastid. Also, two of the enzymes of the type II isoprenoid biosynthesis pathway contain [Fe-S] clusters.[129,130]

Given the multiple membranes of the plastid, it is difficult to imagine how mature [Fe-S] clusters could be imported. Therefore, [Fe-S] clusters may be assembled *in situ.*[131] In light of the already established involvement of the plastid in haem biosynthesis (section 1.2.5.6), this suggests that a major role of the plastid organelle is the metabolism of

iron. The remainder of this chapter describes the current state of knowledge of biological [Fe-S] cluster assembly.

1.3. IRON-SULPHUR CLUSTERS.

1.3.1. [Fe-S] clusters in biology.

1.3.1.1. Discovery of [Fe-S] clusters.

As early as 1958, studies of the mitochondrion conducted by Crane[132] suggested the existence of a novel class of non-haem iron proteins; by 1960, Beinert and Sands[133,134] had expanded these data using electron paramagnetic resonance spectroscopy. In the following decade, [Fe-S] clusters were conclusively identified in biological material using various spectroscopic techniques (e.g., Mössbauer), interpreted according to a new model of electron spin (historical review by Beinert[135]).

1.3.1.2. Structure and function of [Fe-S] clusters.

[Fe-S] clusters exhibit great diversity of structure and function. As a structural approximation, iron is in a tetrahedral configuration with sulphur. Two of the simplest clusters are rhombic [2Fe-2S] and cubane [4Fe-4S]; conversions between clusters readily occur. In [2Fe-2S] clusters, the Fe-S bond measures *ca.* 0.22 nm and clusters adopt a butterfly shape. Attachment to proteins is predominantly through the thiol groups of cysteine residues. According to protein environment, a wide range of reduction potentials (+400 to −700 mV) are available; however, the majority are highly negative, a property exploited in redox catalysis.[136] [Fe-S] clusters are also implicated in non-redox processes, notably, binding and activation of substrates; stabilisation of protein structure; and oxygen sensing.[128]

1.3.1.3. Discovery of the mechanisms of biological [Fe-S] cluster assembly.

Pioneering studies of the nitrogen fixation (NIF) system of the free-living diazotrophic bacterium *Azotobacter vinelandii* conducted by Dean provided clues to the biological assembly of [Fe-S] clusters. Specifically, it was demonstrated that the [Mo-7Fe-9S] cofactor of the nitrogenase was assembled by the action of a cysteine desulphurase (NifS) and a scaffold protein (NifU).[137] [Fe-S] cluster assembly mechanisms have since been discovered in bacteria and eukaryotes. Three main systems are known, namely, NIF; iron-sulphur cluster assembly (ISC); and sulphur mobilisation (SUF). In bacteria, the three [Fe-S] cluster assembly systems are organised into operons (Table 1.1). In *E. coli,* it is worth noting the existence of a fourth cysteine desulphurase (CSD) encoded alongside a SufE-like protein (YgdK); these proteins might contribute to another pathway of [Fe-S] cluster assembly. Organisms use one or more of these systems to assemble [Fe-S] clusters, as has been exhaustively reviewed.[138-142]

1.3.2. Chemistry of [Fe-S] cluster assembly.

1.3.2.1. [Fe-S] cluster assembly.

The chemistry of biological [Fe-S] cluster assembly remains unclear. In particular, how clusters, which are exquisitely oxygen sensitive, can be assembled under aerobic conditions. Nevertheless, study of low molecular weight structural analogues has provided some clues, indicating that clusters are readily formed by reaction of Fe^{2+}, sulphide and thiol in an aprotic, anaerobic solvent.[143]

TABLE 1.1

IRON-SULPHUR CLUSTER ASSEMBLY SYSTEMS IN BACTERIA.

Protein	GenBank number	Annotation
***suf* operon (*Escherichia coli*)**		
SufA	P77667	Alternative Fe-S assembly scaffold
SufB	P77522	Component of SufBCD complex
SufC	P77499	Soluble ABC ATPase
SufD	P77689	Component of SufBCD complex
SufS	P77444	Cysteine desulphurase
SufE	P76194	Acceptor of sulphur from SufS
***isc* operon (*E. coli*)**		
IscR	AAC75584	Transcriptional regulator
IscS	P39171	Cysteine desulphurase
IscU	P77310	Fe-S assembly scaffold
IscA	AAC75581	Alternative Fe-S assembly scaffold
HscA	P36541	Hsp70 class chaperone
HscB	P36540	DnaJ domain protein
Fdx	AAG57639	Ferredoxin
***nif* operon (*Azotobacter* spp.)**		
Nif-IscA	AAC24474	Alternative Fe-S assembly scaffold
NifU	P05340	Fe-S assembly scaffold
NifS	P05341	Cysteine desulphurase
NifV	P05342	Homocitrate synthase
CysE/SAT	P23145	Serine O-acetyltransferase
Orphans (*E. coli*)		
CSD	Q46925	Cysteine desulphurase
YdgK	P76180	Acceptor of sulphur from CSD?

Biological assembly may occur by a similar route but under different conditions, as illustrated *in vitro* with the $[4Fe\text{-}4S]^{3+}$ cluster of the high potential iron-sulphur protein (HiPIP; GenBank IHKREV) of *Allochromatium vinosum.*[144] Such clusters can be assembled on isolated apoprotein (HiPIPapo) by mixing with Fe^{2+} and sulphide in a reducing environment. Mass spectroscopic and enzymatic data suggest that HiPIPapo is first metalled then sulphured (Eq. 1.1).

$$\text{HiPIPapo} \xrightarrow{Fe^{2+}} \text{HiPIP}[4Fe] \xrightarrow{S^{2-}} \text{HiPIP}[4Fe-4S]^{3+} \qquad \text{(Eq. 1.1)}$$

1.3.2.2. [Fe-S] cluster repair.

[Fe-S] clusters are sensitive to damage. Therefore, it is useful to consider mechanisms of repair. Two repair chemistries may exist, reflecting the two types of damage that occur, namely, loss of Fe and loss of S.

It is worth noting that [Fe-S] clusters in different coordination environments of proteins exhibit markedly different susceptibilities to damage. For example, $[4Fe\text{-}4S]^{2+}$ clusters in dehydratases may be especially susceptible to oxidative damage because only three Fe atoms are coordinated by protein cysteine residues; leaving one Fe atom vulnerable to oxidative attack and disassociation. This scenario is classically exemplified by ammonium persulphate action on the $[4Fe\text{-}4S]^{2+}$ cluster of cardiac mitochondrial aconitase (EC 4.2.1.3) of *Bos taurus* (Artiodactyla: Bovidae).[145] Oxidation was shown to destabilise the cluster, leading to the rapid (15-20 min) loss of one Fe atom forming a somewhat stable $[3Fe\text{-}4S]^{1+}$ cluster; however, if oxidative conditions were prolonged (~ h), the cluster disintegrated completely.

Chemistry described in section 1.3.2.1 would be critical for the regeneration of completely disintegrated clusters. However, *in vivo*, it is difficult to envisage a situation where damage would be allowed to proceed unabated without [Fe-S] cluster repair. Hence, i*n vivo,* oxidative attack would generate damaged clusters that the cell would immediately re-metal. A hypothetical scheme offered by Djaman, *et al.*[146] suggests that re-metalling may be achieved by reduction (Eq. 1.2) followed by insertion of iron (Eq. 1.3).

$$[3Fe-4S]^{1+} + e^{-} \rightarrow [3Fe-4S]^{0} \qquad \text{(Eq. 1.2)}$$

$$[3Fe-4S]^{0} + Fe^{2+} \rightarrow [4Fe-4S]^{2+} \qquad \text{(Eq. 1.3)}$$

Parsimony dictates that such a mechanism would require only two catalytic participants, namely, an electron donor, and an iron donor. Actually, oxidatively damaged clusters of the aconitase could be re-metalled *in vitro* by mixing with dithiothreitol (DTT) under anaerobic conditions; Fe^{2+} was supplied endogenously by pilfering from a subpopulation of [Fe-S] clusters.[145] Hence, *in vitro*, repair is spontaneous. However, in *E. coli*, uptake of iron from the extracellular milieu and contributions from the iron storage proteins bacterioferritin and ferritin are implicated in the re-metalling of $[3Fe-4S]^{1+}$ clusters.[147]

The second form of [Fe-S] cluster damage, loss of sulphur, requires re-sulphuration. The paradigm in this case might be NO (g) attack on the [2Fe-2S] clusters of *E. coli* ferredoxin.[148] Biomimetic studies indicated that addition of NO (g) led to the reductive elimination of sulphur from the cluster to form a stable protein-bound dinitrosyl iron

complex [Fe-2NO].[149] Resulphuration *in vitro* was achieved by addition of L-cysteine and a cysteine desulphurase (*E. coli* IscS).[150]

1.3.3. Biology of [Fe-S] cluster assembly.

1.3.3.1. Biological distinction between [Fe-S] cluster assembly and repair.

Chemical distinction between cluster assembly and repair (section 1.3.2) assumes greater significance in consideration of the biology of the process. Specifically, it is important to distinguish between *de novo* cluster assembly; and cluster repair. *De novo* assembly, probably of basic subunits, i.e., [2Fe-2S] and [4Fe-4S], is assumed to occur on a "professional" scaffold protein.[151] Subsequently, pre-assembled [Fe-S] clusters may be passed downstream to target proteins where they can be processed into more complex clusters. Conversely, repair may be affected *in situ* on the target protein, without the involvement of a scaffold.[146] Hence, *de novo* cluster assembly requires a large repertoire of proteins; while for repair, only a subset of the repertoire is needed.

1.3.3.2. Minimum protein components for [Fe-S] cluster assembly.

What constitutes the complete repertoire of proteins for [Fe-S] cluster assembly is inadequately described, but is likely to be extremely complex, involving many proteins and chaperones.[152] Chemical evidence (section 1.3.2.1) implies that *de novo* assembly centres on a three-way interaction between a scaffold protein, an iron donor and a sulphur donor (in a reducing environment); for cluster repair, the scaffold may be omitted. In all cases, energy must also be required.[153] Considerable information about

the process has been derived from the study of the ISC systems of *E. coli* and *Saccharomyces cerevisiae* (Ascomycota).

1.3.3.3. Scaffold proteins.

De novo assembly probably requires a scaffold protein to physically support the formation of [Fe-S] clusters prior to transfer to the target protein. Indeed, as the scaffold must probably interact with numerous different downstream targets, it seems reasonable *a priori* to assume the existence of scaffold proteins bearing different interaction specificities. However, only two types of scaffold protein (A and U) are known. Examples of type A scaffolds include Nif-IscA, IscA and SufA; while examples of type U scaffolds include NifU and IscU. IscA and IscU exhibit structural similarities.[154-156]

1.3.3.4. Iron donors.

An iron donor must be involved, given the cellular toxicity of free iron. Specifically, Fe^{2+} contributes to oxidative stress via Fenton chemistry (Eq. 1.4).[157]

$$Fe^{2+} + H_2O_2 \rightarrow Fe^{3+} + HO^{\bullet} + HO^{-} \qquad \text{(Eq. 1.4)}$$

A yeast protein called frataxin (YFH1) is essential for [Fe-S] cluster formation in isolated yeast mitochondria.[153] This has lead to the suggestion that frataxin is the iron donor for [Fe-S] cluster assembly in the ISC system. Studies on human frataxin revealed that one protein equivalent could retain 6-8 Fe^{2+} and was able to bind human IscU.[158] Indeed, frataxin was directly implicated in the re-metalling of oxidatively

damaged $[4Fe\text{-}4S]^{2+}$ clusters of aconitase.[159] Additional iron donors associated with the NIF and SUF systems are unknown.

1.3.3.5. Sulphur donors.

The sulphur donor is an enzyme with cysteine desulphurase activity.[160] Such enzymes catalyse the oxidation of cysteine to sulphur and alanine. The free L-cysteine concentration of the cytosol in aerobically grown *E. coli* was estimated to be ~ 0.2 mM;[161] an adequate substrate reserve. Four important enzymes showing cysteine desulphurase activity are: NifS,[162] IscS,[163] CSD[164] and SufS.[165,166] It is worth noting that besides [Fe-S] cluster assembly, these enzymes may take part in other cellular processes. For example, IscS could contribute sulphur to thiolated tRNAs[167] and thiazole synthesis;[168,169] while cysteine desulphurases could be involved in selenium metabolism owing to their deselenating action.[170]

1.3.4. Trend of specialisation among [Fe-S] cluster assembly systems.

1.3.4.1. Copy number of [Fe-S] cluster assembly systems.

The existence of three different systems of [Fe-S] cluster assembly that nevertheless share common features (section 1.3.1.3) suggests divergence of function. To consider this further, it is useful to compare organisms possessing one [Fe-S] cluster assembly system with those organisms possessing two or more [Fe-S] cluster assembly systems.

1.3.4.2. One [Fe-S] cluster assembly system.

Numerous species possess only one system of [Fe-S] cluster assembly. Bacterial examples include *Mycobacterium tuberculosis* and *Thermatoga maritima* (SUF); and *Heliobacter pylori* (NIF).* Eukaryote taxa lacking a plastid also possess only one [Fe-S] cluster assembly system; well-characterised examples include *S. cerevisiae* and humans, both exclusively reliant upon the ISC system.[141] Presumably, in these organisms, the solo system is required to fulfil the entire cellular demand for [Fe-S] clusters. In support of this conclusion, it should be pointed out that the NIF system from the eukaryote *Entamoeba histolytica* (Amoebida: Entamoebidae) is functionally interchangeable with SUF and ISC systems, as evidenced by complementation of a SUF/ISC knockout (KO) strain of *E. coli* under anaerobic conditions.[171] This contrasts markedly with the specialised NIF system of *A. vinelandii* (section 1.3.4.3).

1.3.4.3. Two or more [Fe-S] cluster assembly systems.

A more complex situation arises in organisms possessing two or more [Fe-S] cluster assembly systems, as exemplified by numerous bacteria and plastid-bearing eukaryotes (section 1.3.4.4). For example, *A. vinelandii* retains both ISC and NIF systems. In this organism, the ISC system is presumably tasked with the constitutive supply of [Fe-S] clusters ("housekeeping") to proteins, e.g., aconitase; while the NIF system fulfils the specialised role of supplying clusters to the nitrogenase. It is worth noting that ISC can partially substitute the nitrogenase supply role in cells ablated for NIF;[137] conversely, NIF cannot functionally compensate ISC.[172] Hence, the NIF system of *A. vinelandii* is highly specialised for delivery of [Fe-S] clusters to the nitrogenase.

---

* Johnson, D.C., Dean, D.R., Smith, A.D. and Johnson, M.K. (2005) *Annu. Rev. Biochem.*, in press.

In *E. coli*, possessing ISC and SUF systems, and possibly a third system (CSD/YgdK), the division of labour is more opaque. It is however clear that the systems function independently, e.g., *E. coli* IscU can only make a functional (sulphur exchange) interaction with IscS, not with SufS or CSD.[173]

Genetic and transcriptomic studies provided some clues to the division of labour in *E. coli*. For example, an ISC KO strain of *E. coli* was able to grow, albeit slowly, suggesting that SUF was able to compensate.[174] Therefore, it was concluded that while the ISC system provided the bulk of [Fe-S] clusters, SUF supplied a limited number (at least sufficient for survival against an ISC KO background).

Transcriptomic study of the bacterium under oxidative stress indicated that both ISC and SUF systems upregulated;[175] unsurprising given that oxidative stress damages [Fe-S] clusters. However, evidence for differential function came from the finding that ISC and SUF responses to oxidative stress were regulated by different mechanisms. In *E. coli*, ISC genes responded to IscR (GenBank AAC75584), a $[2Fe\text{-}2S]^{1+}$ cluster-containing transcriptional repressor that probably unblocks expression from the *isc* operon when intracellular [Fe-S] cluster levels fall.[176] Oxidative stress damages [Fe-S] clusters, thereby reducing their intracellular concentration and indirectly stimulating ISC. Conversely, the *suf* operon of *E. coli* was directly regulated by OxyR (GenBank AAC76943), a transcriptional activator that responds to peroxide stress by binding to a promoter region 220 bp upstream (sufAp) from the transcriptional start of sufA, thus inducing SUF.[175,177]

These data immediately suggested that the ISC system fulfilled a constitutive role. Additionally, it is worth noting that the ISC system was capable of repairing [Fe-S] clusters damaged by oxidative stress.[146] In contrast, the SUF system came to the fore during prolonged periods of oxidative stress, possibly as a backup; but was unable to repair oxidatively damaged clusters.[146] Prolonged periods of oxidative stress (section 1.3.2.2) are likely to saturate the capacity of repair mechanisms, thus requiring extensive *de novo* synthesis to restock the cell with [Fe-S] clusters.

Recently, Outten, *et al.*[178] have proposed another model of ISC and SUF regulation in *E. coli*. It was suggested that oxidative stress was merely the proximate cause of induction of the SUF system in *E. coli* while the ultimate cause was iron starvation. The rationale given was that the intracellular labile Fe^{2+} pool becomes dangerous under conditions of oxidative stress, due to classical Fenton chemistry (Eq. 1.4). Consequently, the cell may take measures to lower the intracellular iron concentration, including reduction of uptake from the environment; and sequestration into proteins such as DNA binding protein from starved cells.

This scenario would create an iron-limited environment unfavourable to [Fe-S] cluster assembly via the ISC pathway. Hence, it was proposed that the SUF system mediates [Fe-S] cluster assembly under such conditions. While this iron limitation hypothesis awaits further experimental study, it fits with circumstantial evidence. For example, in *E. chrysanthemi,* SUF genes are regulated by the iron-responsive transcription factor Fur (GenBank AAD01582).[179] Additionally, in bacteria lacking an obvious Fur system, e.g., *Rhizobium leguminosarum* biovar *viciae*,[180] regulation of SUF genes by the rhizobial iron regulator (RirA; GenBank CAC35510) has been reported.[181]

1.3.4.4. Compartmental specialisation in eukaryotes.

In plastid-bearing eukaryotes possessing two [Fe-S] cluster assembly systems per cell, division of labour is perhaps related less to function and more to the imperative of intracellular compartmentalisation. In *A. thaliana*, there is good evidence for the localisation of the SUF system in the plastid (section 1.3.6); and the ISC system in the mitochondrion.[182,183] Iron starvation may be a problem in alkaline soils, due to the poor solubility of iron at high pH.[184] However, low iron did not stimulate transcription of *AtsufACDES*, and even downregulated *AtsufB*.[185]

1.3.5. SUF system in bacteria.

1.3.5.1. The suf *operon.*

Patzer and Hantke[166] coined the term sulphur mobilisation (SUF) and implicated SufS and SufD in the assembly of [2Fe-2S] clusters in *E. coli*. Existence of a *suf* operon (sufABCDES) was subsequently shown; all but sufD were essential for viability against an ISC KO background.[174] Indeed, with the exception of sufD, all members of the operon have been studied. The following discussion dissects the function of each protein.

1.3.5.2. E. coli *SufA.*

SufA of *E. chrysanthemi* has been studied by Ollagnier-de Choudens *et al.*[151,186] using recombinant protein probed with Mössbauer spectroscopy. At the protein sequence level, SufA and IscA cannot be distinguished, immediately suggesting a similar

function. In actuality, both proteins operated by a similar mechanism. Specifically, it was shown that [Fe-S] clusters of $[2Fe\text{-}2S]^{2+}$ and $[4Fe\text{-}4S]^{2+}$ type were assembled upon the SufA scaffold; before transfer to a target protein, as demonstrated by the interaction between SufA and biotin synthase. Critical to the transfer reaction was a conserved group of cysteines. Structural information for *E. coli* SufA is presently being derived by NMR.[187]

1.3.5.3. E. coli *SufB and SufC.*

SufBC were shown to be essential by KO in the cyanobacterium *Synechocystis* sp.[188] Bacterial SufC proteins have been thoroughly examined. For example, SufC was implicated in the protection of [Fe-S] clusters under oxidative stress.[189] Primary amino acid sequence of SufC predicts ATPase activity, attributable to standard ATP-binding cassette (ABC) features (Walker box and C motifs). Indeed, Rangachari, *et al.*[190] detected ATPase activity in recombinant *Thermatoga maritima* SufC (K_m = 45 μM); a similar trend was obtained by Nachin, *et al.*[189] for *E. chrysanthemi* SufC (K_m = 290 μM). Interactions between SufBC[189,190] and D[189] were also demonstrated. Furthermore, in *E. coli*, Outten, *et al.*[191] reported that SufBCD was able to increase the cysteine desulphurase activity of SufS by a factor of 32 over isolated SufS. In *Erwinia chrysanthemi*, SufBCD fractionated with the cytosol, suggesting a cytosolic location[189]. Conversely, immunogold EM of SufBC in *T. maritima* unequivocally indicated a membrane location.[190]

1.3.5.4. E. coli *SufS and SufE.*

SufS enzymes of two related enteric bacteria (*E. coli* and *E. chrysanthemi*) and the cyanobacterium *Synechocystis* sp. have been subjected to extensive study (chapter 4). It is vital to note that SufS may interact with other members of the *suf* operon, forming a multicomponent cysteine desulphurase. For example, the activity of *E. chrysanthemi* SufS against L-cysteine was enhanced 50-fold by binding to SufE;[192] the equivalent figure for the *E. coli* proteins was 8-fold.[191] It was subsequently shown by mass spectrometry of the *E. chrysanthemi* proteins that sulphur generated by the action of SufS was transferred to a conserved cysteine (C51) on SufE.[193] Recent structural information for *E. coli* SufE reveals similarity to IscU.[194]

1.3.6. SUF system in vascular plants.*

1.3.6.1. [Fe-S] cluster assembly in the chloroplast.

Numerous proteins found in chloroplasts of vascular plants contain [Fe-S] clusters, suggesting a high demand for *in situ* assembly. Indeed, cluster assembly has been demonstrated in isolated chloroplasts of *Spinacia oleracea* (Chenopodiaceae).[200,201] It is also worth noting the oxygen rich environment of a photosynthesising chloroplast,[202] hinting at the need for efficient [Fe-S] repair.

The system is likely to be complex, but recent bioinformatic analysis by Xu and Møller[197] indicates that *A. thaliana* chloroplasts contain all the proteins of the bacterial

* Designations for SUF proteins differ between the plant and bacterial taxa. For example, in *A. thaliana*, SufA is AtIsa1;[195] SufB is LAF6,[196] AtABC1[196] or AtNAP1;[185] SufC is AtNAP7;[197] SufD is AtNAP6;[197] and SufS is NFS2[198] or CpNifS.[199] For clarity, I exclusively use the bacterial terminology (SufABCDS).

SUF system. These proteins are encoded by the nuclear genome but are imported to the chloroplast.

1.3.6.2. A. thaliana *SufA/IscA and NifU.*

SufA/IscA (TAIR* locus AT1G10500) was detected in the *A. thaliana* genome and predicted to target the chloroplast *in silico.*[197] Localisation of a SufA-like sequence to the chloroplast has been reported in *A. thaliana* (Léon, Briat and Lobréaux, INRA, Montpellier, France, unpublished observation cited by Touraine, *et al.*[195]). Given the need to interact with different target proteins within the chloroplast (section 1.3.3.3), it seems reasonable to predict the existence of alternative scaffold proteins. Indeed, it is worth noting that NifU-like proteins (NFU1-5) are present in the *A. thaliana* genome. NFU1-3 target the chloroplast while NFU1-2 were able to complement yeast strains defective for ISU1.[182] Furthermore, *A. thaliana* lines disrupted in NFU2 exhibited stunted growth as compared to wild type plants; the mutant phenotype was attributed to photosynthetic defects.[195,203] Together, these data hint at the availability of several alternative scaffolds for cluster assembly in the chloroplast.

1.3.6.3. A. thaliana *SufB*

The properties of AtSufB (TAIR locus AT4G04770) have been examined by Xu, *et al.*[185] It was shown that the protein could complement a SufB KO strain of *E. coli* placed under oxidative stress. Further, homodimerisation and heterodimerisation (with AtSufC) was demonstrated by yeast two-hybrid technology and bimolecular fluorescence complementation in cells of *N. tabacum.* Immunocytochemistry in *A. thaliana* localised

* The Arabidopsis Information Resource (http://www.arabidopsis.org/)

SufB to the chloroplast inner envelope.[196] Heterologously expressed AtSufB possessed ATPase activity (K_m = 36 μM) which could be stimulated 2-fold by Fe^{3+}.

1.3.6.4. A. thaliana *SufC*

AtSufC (TAIR locus AT3G10670) is also likely to be functional, as reported by Xu and Møller.[197] Principally, AtSufC was able to rescue the phenotype of a SufC KO *E. coli* strain. Disruption of AtSufC in *A. thaliana* produced an embryo-lethal phenotype in the homozygote; examination of cell ultrastructure revealed morphological abnormalities in the plastid. Finally, transient transfection of *N. tabacum* indicated plastid targeting.

1.3.6.5. A. thaliana *SufD,*

Disruption of AtSufD (TAIR locus AT1G32500) reduced the number of viable seeds, indicating the gene to be important but not essential.[204] Transient transfection with AtSufD transit peptide revealed plastid targeting.[197] Furthermore, yeast two-hybrid analysis[197] indicated that AtSufD interacted with AtSufC.

1.3.6.6. A. thaliana *SufS and SufE.*

SufS (TAIR locus AT4G26500) is found in *A. thaliana.* It targets the chloroplast and shows similar enzymatic activity against L-cysteine and L-selenocysteine to EcSufS *in vitro.*[198,199] While SufS barely contributes to the stromal proteome (0.06 ± 0.02 %[205]), it is probably essential, as indicated by the finding that SufS KO in *Synechocystis* sp. is lethal.[206] Ye, *et al.*[205] demonstrated that purified AtSufS can participate in [Fe-S] cluster assembly *in vitro.* However, AtSufS specific activity was 50-80 fold higher in

chloroplast stroma (as compared to isolated enzyme *in vitro*); furthermore, a proportion of native AtSufS co-purified at a molecular mass of 600 kDa on a gel filtration column. These data imply binding and stimulation by high molecular weight accessory factors, e.g., AtSufBCDE. It is worth noting that sequence encoding AtSufE (TAIR locus AT4G26500) was identified in the *A. thaliana* genome and predicted to target the chloroplast.[197]

1.4. PROJECT OBJECTIVE AND OVERVIEW.

The objective of this project was to investigate an ORF in *P. falciparum* with homology to SufS of *E. coli* and *A. thaliana*, known as *P. falciparum* SufS (PfSufS). In subsequent chapters, I consider my own studies in which I have used a number of different techniques to examine the properties of PfSufS. Chapter 2 describes initial investigations of PfSufS using bioinformatic and simple *in vitro* techniques. Preliminary experiments on the intracellular targeting of PfSufS are discussed in chapter 3. The major focus of this project was to investigate the enzymatic properties of PfSufS. This required the development of a simple colorimetric cysteine desulphurase assay (chapter 4) and the production of recombinant enzyme (chapter 5).

CHAPTER 2.

INITIAL ANALYSIS OF *PLASMODIUM FALCIPARUM* SUFS.

Genome sequencing has facilitated the search for SUF genes in diverse organisms. This chapter briefly describes the various Apicomplexan genomic data sets available, before discussing both the conservation and divergence of SufS sequences across taxa. Descriptions are provided of the putative SUF system in *P. falciparum* and investigations of *P. falciparum* SufS using *in silico* and laboratory techniques.

2.1. INTRODUCTION.

2.1.1. Survey of Apicomplexan genomes.

2.1.1.1. Apicomplexan nuclear genomes.

Numerous Apicomplexan nuclear genomes have been, or are being, sequenced. Whole genome information is currently available for *P. falciparum* 3D7[45] and *P. yoelii yoelii* 17XNL.[207] The *P. falciparum* genome is further discussed in section 2.1.2. Incomplete genomic information has been deposited for *P. vivax* and several non-human models, notably simian (*P. knowlesi*) and rodent malarias (*P. berghei* and *P. chabaudi chabaudi*). All these data are available via the Internet in PlasmoDB.[208] Sequence data also are available for other Apicomplexan organisms, with a bias towards the major human and veterinary pathogens, e.g., the complete genomes of *Cryptosporidium hominis* TU502[209] and *C. parvum* IOWA;[210] alongside patchy coverage of *Eimeria tenella*, *Perkinsus marinus*, *Theileria parva* and *T. gondii*. Available via the Internet are genomic data for *Cryptosporidium* spp.[211] and *T. gondii*.[212]

2.1.1.2. Apicomplexan organellar genomes.

Organellar genome sequences are available for *P. falciparum*,[53,213] *T. gondii*[54] and *E. tenella*.[56]

2.1.2. *Plasmodium falciparum* genome.

2.1.2.1. Holotypic strain of P. falciparum.

A common laboratory strain (3D7) of *P. falciparum* was selected as the holotype for the genome sequencing project.[45] However, there is considerable historical distance between the holotype and parasites directly isolated from infected patients. Specifically, 3D7 was cloned by limiting dilution[214] from *P. falciparum* NF54, a strain isolated from a baggage handler at Schipol Airport (Amsterdam, The Netherlands).[215] In the intervening years, the cell line has been repeatedly passaged *in vitro*. Hence, it has been relieved of immune pressure in the human host, and out-crossing in the mosquito.

2.1.2.2. Structure of the nuclear genome.

The nuclear genome of *P. falciparum* 3D7 is composed of 22.8 megabases (Mb) of DNA packaged into fourteen chromosomes.[45] The adenine and thymine (AT) content of the DNA is high (80.6 %[45]), a feature that poses a major challenge to sequence analysis and molecular biology.

2.1.2.3. Structure of the organellar genomes.

DNA is additionally contained in the mitochondrion and plastid organelle of *P. falciparum*. The mitochondrial genome is 6 kilobases (kb) long[213] while the plastid genome is 35 kb long.[53] Both organellar genomes contain a minimal set of genes.

2.1.2.4. Prediction of open reading frames in the nuclear genome.

Prediction of ORFs in the nuclear genome of *P. falciparum* has generally been done automatically using algorithms such as GlimmerM and Genefinder; the problems of such *ab initio* gene prediction are well-known[216] and are especially acute in AT-rich DNA as intron boundaries are difficult to identify. Intracellular localisation of proteins has also been predicted using automatic methods, e.g., PlasmoAP and PlasMit (chapter 3).

2.1.2.5. Transcriptomic and proteomic analysis.

A number of transcriptomic and proteomic analyses of *Plasmodium* spp. have been conducted. Of particular note are studies of the *P. falciparum* erythrocytic[18,83,217-219] and sporozoite[83] stages. One EST study is available for the liver stages of *P. yoelii*.[88] Also published are studies of the *P. berghei* erythrocytic[89] and sporozoite[220] stages.

Environmental modulation of transcription has been investigated. For example, a serial analysis of gene expression of the response of erythrocytic stages of *P. falciparum* to the antimalarial drug chloroquine.[221] Proteomic analyses have been conducted for the erythrocytic stages of *P. falciparum*;[86,87] while a study of the mosquito stages of *P. berghei* recently became available.[89] Unsurprisingly, meta-analysis[222] has revealed significant inconsistencies between studies.

2.1.3. Cautionary remarks.

While genomic, transcriptomic and proteomic data facilitate investigation of the biology of Apicomplexan organisms, it is essential to be cautious. In particular, the AT bias of *Plasmodium* spp. DNA makes gene prediction unreliable. Data from transcriptomic studies must also be interpreted with care. Finally, it is easy to challenge the validity of *P. falciparum* 3D7 as a model for clinical malaria.

2.1.4. Chapter objective and overview.

The objective of this chapter is to set the scene for subsequent data by reporting my initial studies of PfSufS. Results in this chapter derive from bioinformatic analysis of published genomic and transcriptomic information and simple *in vitro* experiments. Following the Materials and Methods, the results section begins with a discussion of [Fe-S] cluster assembly in *P. falciparum* (emphasis on the putative SUF system). Subsequently, the focus shifts to SufS. I provide a definition of SufS in terms of amino acid sequence motifs. I report a survey of genomes encoding SufS among the bacterial, protozoan, algal and vascular plant taxa. Based on computational studies, a description of PfSufS at nucleotide, amino acid and structural levels is presented. This is followed by *in vitro* evidence for the chromosomal location and transcription of PfsufS. Finally, I describe my investigations of the transcriptional control of *PfsufS* under conditions of oxidative stress and iron limitation.

2.2. MATERIALS AND METHODS.

2.2.1. Computational methods.

2.2.1.1. Nucleic acid sequences.

Rates of synonymous (dS) and non-synonymous (dN) mutation were estimated according to the method of Goldman and Yang[245] implemented in the codeml algorithm in the PAML software.[246*] Conserved motifs were identified with MEME.[247] Quality of PCR primers was checked with the Amplify software. DNA sequences were analysed using the ABI Prism Sequencing Analysis software (Perkin Elmer) running on an Apple Power Mac G4 and assembled using the SeqMan software (DNASTAR).

2.2.1.2. Protein sequences.

SufS sequences were retrieved from genome databases using the TBLASTN algorithm with EcSufS (GenBank P77444) or AtSufS (TAIR locus AT4G26500) as probes. Amino acid sequences were aligned using T-COFFEE[248†] or Clustal W[249‡] and refined by eye. Phylogenetic trees were constructed from protein sequence alignments using the PHYLIP software (J. Felsenstein, University of Washington, USA[¶]) running on an Apple iBook. Trees were constructed by: (1) neighbour-joining (FITCH); (2) parsimony (PROTPARS); and (3) maximum likelihood (PROTML) according to the JTT probability model.[250] Consensus trees were abstracted using CONSENSUS and drawn with DRAWTREE. Bootstrapping (100 pseudoreplicates) was done with SEQBOOT.

---

* http://abacus.gene.ucl.ac.uk/software/paml.html
† http://www.ch.embnet.org/software/TCoffee.html
‡ http://www.ebi.ac.uk/clustalw/
¶ http://evolution.gs.washington.edu/phylip.html

2.2.1.3. Protein properties and structure.

Kyte-Doolittle hydropathy plots were constructed with an Internet-based implementation.[233]* The ExPASy ProtParam tool† was used to predict the molecular weights of proteins from primary sequence. Molecular modelling was conducted by Dr. J. W. Saldanha (NIMR). A homology model of PfSufS was produced via threading using the 2.0 Å X-ray crystal structure of EcSufS (PDB 1JF9[251]) as a template. Molecular models were manipulated with the RasMol software.

2.2.2. *Plasmodium falciparum* red blood cell stage culture.

P. falciparum 3D7 and C10 erythrocytic stages were cultured *in vitro* by standard methods[252] in leukocyte-depleted whole blood obtained from the National Blood Service (UK). C10 is a cloned line derived from an artificial mixture of BW (The Gambia) and Palo Alto (Uganda) isolates.[253]

2.2.3. Molecular biology techniques.

P. falciparum 3D7 genomic DNA was a gift from Dr. B. Clough (NIMR). Total RNA was isolated from *P. falciparum* 3D7 mixed erythrocytic stage parasites (parasitaemia *ca.* 10 %) essentially as described[254] but using the TRIzol reagent (Invitrogen). DNA was removed from RNA preparations by treatment with TURBO DNase (Ambion) for 30 min at 37 °C. Quality of isolated RNA was verified on agarose gel by inspection of

---

*http://occawlonline.pearsoned.com/bookbind/pubbooks/bc_mcampbell_genomics_1/medialib/activities/kd/kyte-doolittle.htm

† http://www.expasy.org/tools/protparam.html

28S and 18S ribosomal bands. Copy DNA was made using the RETROscript kit (Ambion) while PCR was done with AmpliTaq (Perkin Elmer) or the LA PCR kit version 2.1 (Takara). Primers were obtained from Eurogentec or Sigma-Genosys. Primer sequences are indicated in Table 2.1. PCR was done on a Biometra T3 Thermocycler. PCR products were purified with the QIAquick or QIASpin kits (Qiagen). To avoid nicking, excision of DNA from agarose gels was guided by blue light (Dark Reader Transilluminator, Clare Chemical Research). Restriction enzymes were obtained from Roche or NEB. Vector inserts were confirmed by restriction digest and Sanger sequencing using the BigDye Terminator v1.1 Cycle Sequencing chemistry (Applied Biosystems) resolved on an ABI Prism 377 DNA sequencer (Perkin Elmer). Plasmids were maintained in INVαF′ *E. coli* cells (Invitrogen). All other molecular biology procedures were done using standard techniques.[255]

2.2.4. Recovery of genes by PCR.

A 690 bp fragment of *P. falciparum* sufB (PfsufB[690]; GenBank CAA64569) was amplified with primers 2.1 and 2.2. Thermocycling conditions were: denaturation at 95 °C for 3 min; followed by 34 cycles of 95 °C for 30 s; 45 °C for 30 s; and 68 °C for 2 min; followed by 72 °C for 10 min.

PfsufS (PlasmoDB PF07_0068) predicted to encode the mature enzyme (D62-K546; PfSufS[M]) was recovered from cDNA using two overlapping sets of primers in two separate PCR reactions. The 5′ region of PfsufS[M] was recovered with primers 2.3 and 2.4; the 3′ region was recovered with primers 2.5 and 2.6.

TABLE 2.1.

PCR PRIMERS USED IN CHAPTER 2.

ID	DNA sequence
2.1	5′TAAATTTATATTTAATAAGACAAGGATTA
2.2	3′TAACCATATCTTTTACTATTATTTCTACTATAGC
2.3	5′GAGCTCGATTATTTTAAAAATGTAAGAGAACAT
2.4	3′CATTAGTTATTATATTACTACCAT
2.5	5′GTGCATCACTAGGTACAGGT
2.6	3′CTCGAGTTTTTCATTTTTCATTTCATTTAACAT
2.7	5′AAAATGAGTGAGTTTTATAAAAAGG
2.8	3′TTTTTCATTTTTCATTTCATTTAAC
2.9	5′ATGGGAGAAGAAGATGTTCAAGCTTTAG
2.10	3′GAAACATTTTCTGTGGACAATACTTGGTCC

Critically, primers 2.3 and 2.4 amplified sequence across the intron boundary, thus allowing the detection of contaminating genomic DNA. Thermocycling conditions for both reactions were: denaturation at 94 °C for 1 min; followed by 30 cycles of 98 °C for 20 s; 45 °C for 30 s; and 72 °C for 1 min.

PfsufS exon 2 (PfsufS[E2]) was amplified from genomic DNA with primers 2.7 and 2.8. Thermocycling conditions were: denaturation at 95 °C for 3 min; followed by 34 cycles of 95 °C for 30 s; 45 °C for 30 s; and 68 °C for 1.5 min; followed by 72 °C for 10 min. The fragment was inserted into the pCRII TA cloning vector (Invitrogen) between Sp6 and T7 promoters. This construct was known as pCRII:PfsufS[E2].

P. falciparum actinI[256] (PfactinI; GenBank P10988) was recovered from genomic DNA by PCR using primers 2.9 and 2.10. Thermocycling conditions were: denaturation at 95 °C for 3 min; followed by 34 cycles of 95 °C for 30 s; 64 °C for 2.5 min; followed by 72 °C for 10 min. Amplicons were cloned into pCRII behind the Sp6 promoter, giving pCRII:PfactinI. Sequence was confirmed by Sanger sequencing.

2.2.5. Chromosomal localisation of PfsufS.

2.2.5.1. Pulsed field gel electrophoresis and blotting.

Chromosomes were separated by pulsed field gel electrophoresis (PFGE). *P. falciparum* 3D7 and C10 red blood cell stage parasites were embedded in blocks of low melting point (LMP) agarose (Seaplaque GCG) and extracted by treatment with 4 % sodium lauryl sarcosinate in 250 $\mu g.mL^{-1}$ proteinase K[255] by Mr. P. W. Moore and Dr. D. H. Williamson (NIMR). Blocks were subjected to PFGE through 1 % chromosomal grade

agarose (BRL) using a custom-built contour clamped homogenous electric field (CHEF) apparatus, as described.[91] Yeast genomic DNA was run as a control. The gel was Southern blotted to Hybond N+ membrane (Amersham) as described.[255]

2.2.5.2. Synthesis of radiolabelled probe.

PfsufS[E2] was excised from pCRII:PfsufS[E2] by *EcoRI* digest and run into LMP agarose. PfsufS[E2] DNA was eluted from LMP agarose into redistilled water by heating at 65 °C for 5 min, followed by boiling for 3 min. Eluted DNA (25 ng per reaction) was radiolabelled with [α^{32}P]adenosine triphosphate (Amersham) using the Prime-it II kit (Stratagene). The hybridization probe was purified through a ProbeQuant G-50 micro column (Amersham).

2.2.5.3. Probing CHEF blot.

CHEF blots were immersed in hybridisation buffer (0.5 % dried milk, 36 mM NaCl, 20 mM sodium phosphate, 2 mM EDTA, 1 % SDS, 250 μg.mL^{-1} salmon sperm DNA, 10 % dextran sulphate in redistilled water) and incubated at 65 °C for 2 h in a Hybaid Micro 4 hybridisation rotisserie. Radio-labelled PfsufS[E2] probe (50 μL) was added and the mix incubated for a further 12 h. Blots were washed at 65 °C sequentially in 2 ×, 1 × and 0.1 × SSPE diluted from 20 × SSPE (140.26 g NaCl, 88.23 g tri-sodium citrate, 35.38 g KH_2PO_4, 7.44 g EDTA in 1 L redistilled water). The blot was autoradiographed.

2.2.6. Response of parasites to iron limitation and oxidative stress.

2.2.6.1. mRNA assay.

Asynchronous *P. falciparum* 3D7 erythrocytic stages (25 mL; 13 % parasitaemia; 1 % haematocrit) were treated with either 2,2′-bipyridyl (BIP) or H_2O_2. BIP (Sigma) was dissolved in DMSO and added at concentrations of 10 or 0 mM. H_2O_2 (Sigma) was added at concentrations from 0-100000 parts per million (ppm). Parasites were incubated with BIP or H_2O_2 for 3 h at 37 °C under 5 % O_2, 5 % CO_2 in N_2. Total RNA was immediately isolated.

2.2.6.2. RNase protection assay.

RNAse protection was used to quantify the level of *PfsufS* mRNA under different BIP and H_2O_2 treatments. PfsufS[E2] was derived from pCRII:PfsufS[E2] by blunt-end digestion with *DraI*. PfactinI was used as a loading control and derived from pCRII:PfactinI by blunt-end digestion with *SspI*. In order to generate radio-labelled RNA probes, digested vectors were transcribed *in vitro* using Sp6 polymerase and [α^{32}P]uracil triphosphate (Amersham), gel purified and eluted in Ambion elution buffer at 37 °C for 12 h. RNA probes were quantified by counting on a Beckman LS6000IC spectrometer.

RNAse protection assay was performed with the RPAIII kit (Ambion), essentially as described by the manufacturer. In a typical reaction, 100 ng total RNA was mixed with 50000 counts per minute of RNA probe and hybridised at 42 °C for 48 h. The hybridisation reaction was digested at 37 °C for 30-120 min with RNAse H/T1

(Ambion) to remove single stranded RNA. The reaction was separated on 6 % acrylamide, 8 M urea gel and autoradiographed. Transcripts bound to RNA probe were protected from digestion and appeared as bands on the autoradiograph.

2.3. RESULTS AND DISCUSSION.

2.3.1. ISC system in the mitochondrion.

Putative genes specifying components of the ISC system have been identified in the *P. falciparum* genome. *P. falciparum* IscS (PfIscS; PlasmoDB MAL7P1.150) has been shown to target the mitochondrion *in vivo*.[223] N-terminal sequences of *C. parvum* IscS and IscU address GFP to the mitochondrion of *S. cerevisiae*.[224] Thus, mirroring the situation in other eukaryotic cells, the ISC system of Apicomplexa is probably mitochondrial.

2.3.2. SUF system in the plastid.

2.3.2.1. Database mining for SUF genes.

It is probable that all members of the SUF system are encoded by *P. falciparum* (Table 2.2; Ellis, *et al.*[225]; Sato, Clough, Rangachari and Wilson, NIMR, unpublished observations). As SufA and IscA cannot be distinguished by amino acid sequence (chapter 1), there are at least two candidates for PfSufA (PfSufA1-2). PfNFU may be included in the SUF system (loosely defined), as it could provide an alternative scaffold for [Fe-S] cluster assembly. It would not be surprising to find multiple scaffolds given the need to interact with, and load [Fe-S] clusters onto, multiple downstream targets. Plastid-localised ferredoxin may act as reductant for [Fe-S] cluster repair.

TABLE 2.2.

SUF SYSTEM PROTEINS IN THE GENOME OF *PLASMODIUM FALCIPARUM.*

Accession number for PfSufB refers to GenBank. All other accession numbers refer to PlasmoDB.

Protein	Accession	Genomic location	Putative function
PfSufA1	PFB0320c	Chr 2	[Fe-S] cluster assembly scaffold
PfSufA2	PFE1135w	Chr 5	[Fe-S] cluster assembly scaffold
PfNFU	PFI1050c	Chr 9	[Fe-S] cluster assembly scaffold
PfSufB	CAA64569	Plastid genome	Component of SufBCD complex
PfSufC	PF14_0133	Chr 14	Soluble ABC ATPase
PfSufD	PF11_0044	Chr 11	Component of SufBCD complex
PfSufS	PF07_0068	Chr 7	Cysteine desulphurase
PfSufE	PFB0270w	Chr 2	Acceptor of sulphur from PfSufS

2.3.2.2. Evidence for plastid location of SUF system.

Circumstantial evidence for the plastid localisation of the SUF system in *P. falciparum* can be marshalled from several directions. Firstly, PfSufB is encoded by the 35 kb plastid genome; transcription of the gene has been observed (Fig. 2.1).[18] Secondly, there is bioinformatic evidence for localisation of PfSufS to the plastid organelle (chapter 3). Given the unreliability of gene models, bioinformatic tests of intracellular localisation are generally uninformative for other members of the SUF system (data not shown). However, PfSufC and PfNFU have been localised to the plastid in *P. falciparum* using a fluorescent reporter fusion (Sato, Thakrar, Clough and Wilson, NIMR, unpublished observations). Thirdly, in Apicomplexan cells lacking a plastid (e.g., *C. parvum*), the SUF system is absent. Finally, compelling analogy can be made to the chloroplastic SUF system of *A. thaliana.*

2.3.2.3. Transcription of SUF genes.

Although the SUF genes are scattered around the genome of *P. falciparum*, their expression might be coordinated at transcriptional and translational levels (based upon the ancestral prokaryotic operon model). Transcriptional co-regulation of proteins that physically interact (e.g., PfSufSE) can be tight.[226] Furthermore, *P. falciparum* genes in the same biochemical pathway are believed to be co-regulated at the transcriptional level.[227]

However, no obvious correlation could be observed in published microarray data for the candidate SUF genes (Fig. 2.2).

FIGURE 2.1.

***PLASMODIUM FALCIPARUM* SUFB.**

A fragment of PfsufB (PfsufB[690]) was recovered by PCR from cDNA (R) and genomic DNA (G) extracted from *P. falciparum* 3D7. Prior to RT-PCR, genomic DNA contamination was removed from RNA preparations by treatment with TURBO DNAse. Control PCR lacking nucleic acid is presented (-). Other abbreviations are as follows. bp: base pair; M: 123 bp ladder (Sigma).

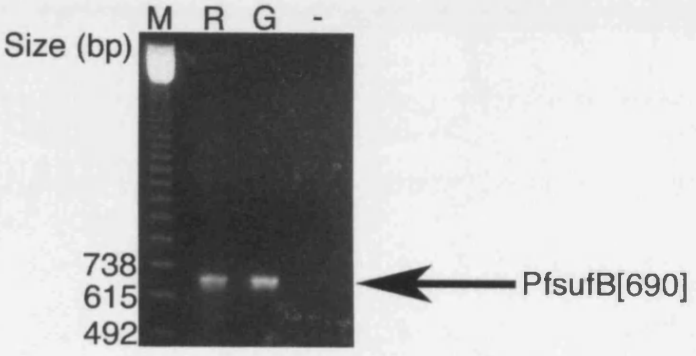
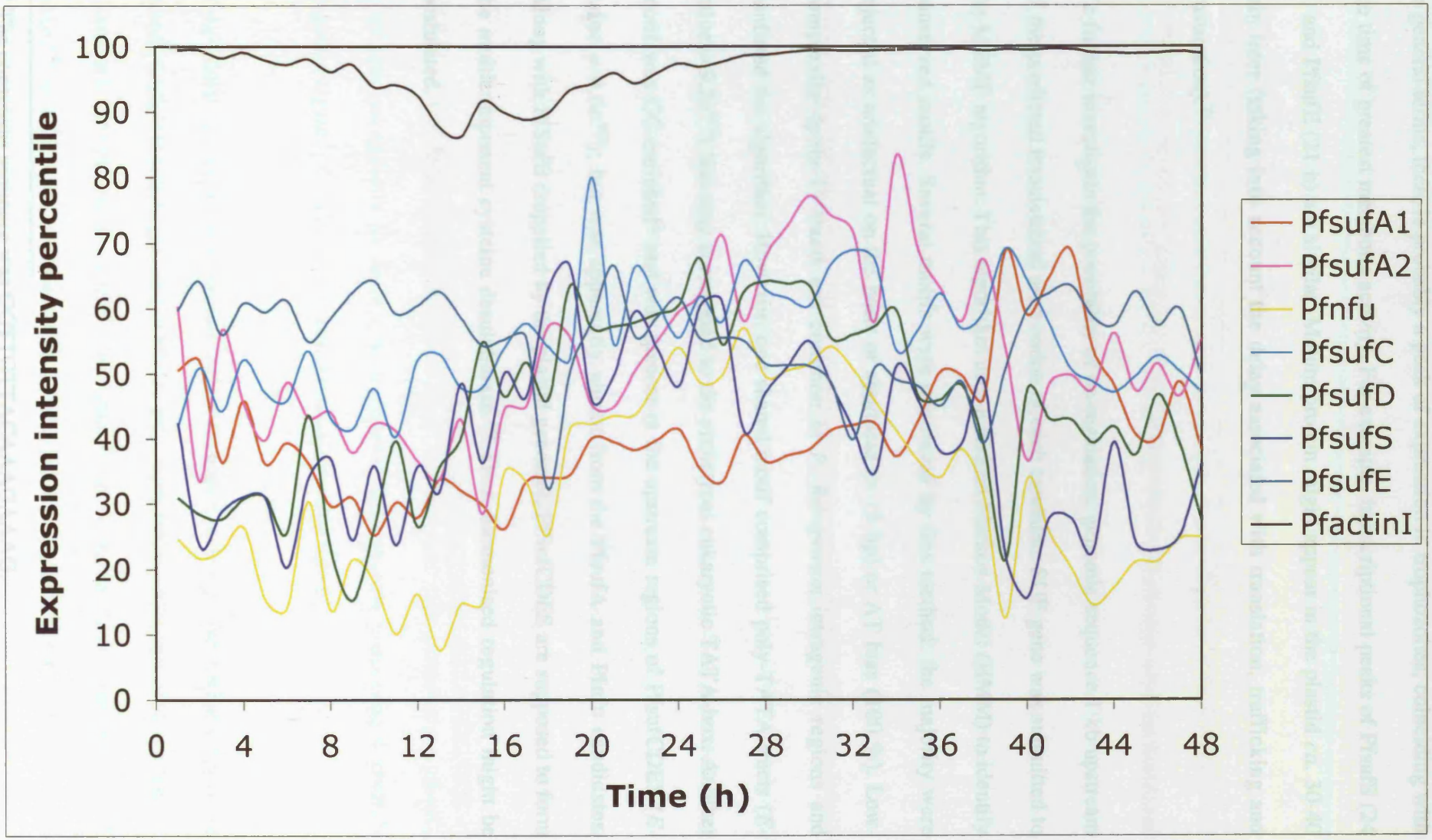

M R G -
Size (bp)
738
615
492
PfsufB[690]

FIGURE 2.2.

TRANSCRIPTIONAL TIME COURSE OF *PLASMODIUM FALCIPARUM* SUF.

Transcriptomic information for the SUF system genes in erythrocytic stages of *P. falciparum*[18] was downloaded from PlasmoDB and plotted as a line graph. The ordinate is reported as a percentage of the expression level of the most abundant transcript at that time point. The transcriptional profile of PfactinI is given for comparison.


Expression intensity percentile
100
90
80
70
60
50
40
30
20
10
0
0
4
8
12
16
20
24
28
32
36
40
44
48
Time (h)
PfsufA1
PfsufA2
Pfnfu
PfsufC
PfsufD
PfsufS
PfsufE
PfactinI

In general terms, there is probably a peak of expression in trophozoites, coinciding with the time of greatest metabolic activity. For example, transcriptional peaks of PfsufS (24 h) and PfsufE (21 h) are similar. Mature protein might appear in the plastid *ca.* 30-40 min later (taking into account the delay associated with translation, trafficking and maturation).[77]

To further investigate the possibility of co-regulation, genomic sequence 1 kb upstream of the predicted translational start codon of each candidate SUF gene was submitted to the MEME algorithm. This algorithm uses a Hidden Markov Model (HMM) to identify conserved motifs. Several motifs were uncovered by this method: the majority were rejected as artefactual on the basis of shortness (> 15 bp) or AT bias (100 %). Low-complexity (poly-T) tracts are common in *P. falciparum* intragenic regions and confused the algorithm. However, one shared motif comprised poly-TATA tracts (*E*-value = $9.2e^{-066}$) that may correspond to the archetypal eukaryotic TATA-box. Another motif was GC-enriched* and was present in the upstream regions of PfsufCDES (*E*-value = $4.6e^{-002}$); but was apparently absent from the PfsufA and Pfnfu candidates. Along with PfSufB (supplied by the plastid genome), PfSufCDES are supposed to form the multicomponent cysteine desulphurase.[191] Thus, harmonised regulation might be postulated.

* The consensus sequence was GGTTGTTACAAAGAAAG.

2.3.3. Definition of SufS.

2.3.3.1. Central importance of SufS in [Fe-S] cluster assembly.

Cysteine desulphurases (EC 2.8.1.7) are a phylogenetically ubiquitous class of enzymes that hydrolyse L-cysteine to L-alanine and sulphur. They are all dependent on pyridoxal 5′-phosphate (PLP) and fall into the aspartate aminotransferase (AAT) category. Cysteine desulphurases may be considered the hallmark of [Fe-S] cluster assembly. Therefore, for the remainder of this thesis, I focus on SufS as a surrogate for the entire SUF system.

2.3.3.2. Definition of SufS at the level of amino acid sequence.

In order to interrogate genomic databases for SufS, it is vital to have an exact definition of the enzyme at the level of primary amino acid sequence. Noting the differences in active site motif among the cysteine desulphurases, Mihara and Esaki[160] subdivided them into class I and II. According to this system, class I enzymes contain variants of the SSGSACTS motif, while class II enzymes contain the GHHCA motif. By these criteria, IscS is class I; while SufS is class II. In eukaryotic cells possessing a plastid, SufS may secondarily be defined as a cysteine desulphurase possessing a plastid-targeting signal (c.f., IscS, with a mitochondrial targeting signal).

If the GHHCA criterion is weakened, highly divergent taxa, e.g., the Archaea, appear to encode SufS. Most divergent is a NifS-like gene from *Halobacterium salinarum* NRC-1 (GenBank NP_281065), an obligately halophilic Archean, possessing a GDHCT motif.[228] In this organism, two deviations from the expected Mihara-type motif are

noted (H → D; A → T). However, it is evident that such changes are minor as both sets of amino acids exhibit similar properties: H and D are charged and hydrophilic; while A and T are small and hydrophobic. It is worth noting that the active site motif of *T. gondii* SufS (TgSufS) exhibits a departure from the expected (GHHCC). However, once again, the change is trivial given that A and C are both small and hydrophobic amino acids. An A → C change is also seen in the last position of the active site motif of other SufS proteins, e.g., in the cyanobacterium, *Anabaena* sp. Taken together, these data suggest considerable plasticity in the active site sequence, provided that there is conservation of amino acid properties. In contrast, the chemistry of cysteine desulphuration demands immutability of the PLP-binding K and active site C (chapter 4). In conclusion, I classified protein sequence as SufS if it contained both a PLP-binding motif (SXHK) and a modification of the Mihara-type active site motif (GXHCX).

2.3.4. Survey of SufS-like sequences.

2.3.4.1. SufS-like sequences are ubiquitous.

SufS-like sequences are present as a single copy in the genomes of organisms from all kingdoms except Metazoa and Fungi: these two groups apparently rely on the ISC pathway (chapter 1). SufS sequences from different taxa were aligned (Fig. 2.3). Inspection of the sequence alignment revealed many short regions of conservation.

Representative sequences between the two anchor motifs (SXHK and GXHCX) were chosen for construction of a phylogenetic tree. For this purpose, large insertions present in the Apicomplexan sequences were also excised.

FIGURE 2.3.

MULTIPLE SEQUENCE ALIGNMENT OF SUFS PROTEIN SEQUENCES.

SufS sequences are found in all taxa except fungi and Metazoa. Essential catalytic residues are conserved across taxa, e.g., the PLP-binding (yellow) and active site (red) motifs. PfSufS possesses a penta-isoleucine tract and an insertion of 51 amino acids (both marked in blue). For this alignment, SufS sequences were chosen to exemplify taxonomic ubiquity. Prokaryotic sequences are taken from the Archaea (*Halobacterium salinarum* and *Pyrococcus furiosus*), enterobacteria (*Escherichia coli* and *Erwinia chrysanthemi*), α-proteobacteria (*Legionella pneumophila*), β-proteobacteria (*Nitrosomonas europaea*), δ-proteobacteria (*Bdellovibrio bacteriovorus*), cyanobacteria (*Anabaena* sp.), high GC Gram+ bacteria (*Mycobacterium tuberculosis*) and spore-forming bacteria (*Bacillus anthracis*). Eukaryotic sequences are from an Apicomplexan parasite (*P. falciparum*), a red alga (*Cyanidioschyzon merolae*), a monocotyledonous plant (*Oryza sativa*) and a dicotyledonous plant (*Arabidopsis thaliana*). Some of the sequences have been truncated at the N-terminal. Abbreviations are as follows. HSALI: *H. salinarum*; PHORI: *P. furiosus*; ECOLI: *E. coli*; ECHRY: *E. chrysanthemi*; NEURO: *N. europaea*; LPNEU: *L. pneumophila*; BBACT: *B. bacteriovorus*; ANABA: *Anabaena* sp.; MTUBE: *M. tuberculosis*; BANTH: *B. anthracis*; PFALC: *P. falciparum*; CMERO: *C. merolae*; OSATI: *O. sativa*; ATHAL: *A. thaliana.*

```
CLUSTAL FORMAT for T-COFFEE Version_1.41, CPU=58.23 sec, SCORE=68, Nseq=14, Len=641

HSALI            --------MEATEQGPIDVEAVRADFPILEREVAGGEDLVYLDNAATSHTPEPVVDAIAD
PHORI            ---------------MKIPDDVRKDIPLTQE-------VIYFDNTATSLTPKPVIEAMDE
ECOLI            --------------MIFSVDKVRADFPVLSREVN-GLPLAYLDSAASAQKPSQVIDAEAE
ECHRY            ---------MNYPVEHYPIDRVRADFPILQQSVN-GQPLAYLDSAASAQKPLAVIDRERD
NEURO            PERHDRAAQIPAAHPPFDIHAIRRDFPILQEHVN-GHSLVWLDNAATTQKPQAVIDRLTY
LPNEU            ------MSTNPSLMTSFDVNKIRKDFPVLHQKIN-EYDLVYFDNAATTQKPKAVIDAIAQ
BBACT            -----------MSNLDDIFASVRSQFPALTQKVH-GKNLVYLDSGATTLKPQSVVDRIAH
ANABA            MT---------YTPTKTLADKVRADFPILHQEVN-GKTLVYLDNAATSQKPLFVLNALRD
MTUBE            ---------MTASVNSLDLAAIRADFPILKRIMRGGNPLAYLDSGATSQRPLQVLDAERE
BANTH            ----------------MNIHEIRKQFPILDQKVN-GKQLVYFDSAATSQKPIQVIETLER
PFALC            --------------IIDYFKNVREHFPFFKEN----KSLIYFDSAATTHKPSCVIEKMSE
CMERO            VEAPGH-----RRNRPAIAALDKDSFPILERTTTEGKQLVYLDSAATSQKPLQVIEALSH
OSATI            RE---------AESAASLGDLTRVDFPILDQEFD-GSKLVYFDNGATSQKPSCVMKTLDE
ATHAL            ST---------DSESVSLGHRVRKDFRILHQEVN-GSKLVYLDSAATSQKPAAVLDALQN
                                       :  :    .       : ::*. *::  *  *:.

HSALI            YYRRYNSNVHRGLHELSQEASVAYEDAHDKLAAFVGGEDREEMVFTKNTTEAENLVAFAW
PHORI            YYLKYRANVHRGVHRLSQMATQKYEESRKVVADFINAK-FEEIVFTKNTSESLNLVALGL
ECOLI            FYRHGYAAVHRGIHTLSAQATEKMENVRKRASLFINARSAEELVFVRGTTEGINLVANSW
ECHRY            FYLHEYAAVHRGIHTLSARATSAMEEVRAKVATFIHAASAEDIVFVRGTTEAINLVANSY
NEURO            FYTHENSNIHRAAHELAARATDAYEESRNKVRQFLNAASTDEIIFLRGTTEAINLVAQSW
LPNEU            FYEKDNSNVHRGVHALSVRATEMYEAARAKVKRFINARSPRECIFVRGTTEAINLVAQSL
BBACT            FYSYETSNVHRGAHYLGDVATGHFEAARQKVAEFLGARQSEEIIFVRGTTEGVNLVANSW
ANABA            YYEQYNSNVHRGAHTLSAKATDAYEGARDKVAKFINAASRKEIVYTRNASEAINLVAYSW
MTUBE            FLTASNGAVHRGAHQLMEEATDAYEQGRADIALFVGA-DTDELVFTKNATEALNLVSYVL
BANTH            YYKEYNSNVHRGVHTLGTKATDAYEGAREKVRKFINAKSMEEIIFTCGTTTALNTVAASY
PFALC            FYKKENSNIHRGIYKLSHNATNNYEKVRETIKEYINCEKNDNIIFTNGSTYGLNVVCKMM
CMERO            FYRYENANVHRGAHRLSQQATEAFESARAKVARFLNAKKDTEIVFTRNATEAINTVAYAW
OSATI            YYRFYNSNVHRGIHVLSAKATDAYESARTKVANFVNAANSREIVFTRNATEAINLVAYSW
ATHAL            YYEFYNSNVHRGIHYLSAKATDEFELARKKVARFINASDSREIVFTRNATEAINLVAYSW
                 :     . :**. : *   *:   *  :     ::      : ::  .:: . * *.

HSALI            GL--NEL----GPG-DEVVLTQMEHHASLVTWQ-QVADETGAEVKYIPITDD-GHLDMDA
PHORI            E---HTF----KKG-DKIVTTPYEHHSNLLPWQ-RLAKKLGLKLEFIEGDDE-GNLDLSD
ECOLI            GN--SNV----RAG-DNIIISQMEHHANIVPWQ-MLCARVGAELRVIPLNPD-GTLQLET
ECHRY            GR--TAF----QPG-DNLVISEMEHHANIVPWQ-MLAQARGLTLRVLPITDD-GELDMAQ
NEURO            GR--QHI----SAG-DEIIITWLEHHANIVPWQ-MLCNETGARLRVVPVDED-GQVLLEE
LPNEU            VA--PRI----LPD-EEILITHMEHHSNIVPWQ-MVCKKMGCKLQVAPISLN-GEVILEE
BBACT            GL--SNL----KAG-DEILITVMEHHGNIVPWQ-MVAEKVGAKVVAADILDN-GELDLED
ANABA            GM--NNL----QAG-DEIILSVMEHHSNIVPWQ-FVAQKTGAVLKFVELTPE-QTLDMEQ
MTUBE            GD--SRFERAVGPG-DVIVTTELEHHANLIPWQ-ELARRTGATLRWYGVTDD-GRIDLDS
BANTH            GL--ENV----KEG-DEIVISYMEHHSNIIPWQ-QVAKKTGATLKYLPLQPD-GTISIED
PFALC            IE--EII----KKEEDEIYLSYMEHHSNIIPWQEYINKEKKGRIKYVPLNKS-GYINIKK
CMERO            GLLGGNL----TPG-DEIIVSVAEHHANLVPWQ-LVAQRTGARLRHIPLLPQVQALDMDT
OSATI            GM--SNL----KQG-DEIVLTIAEHHSAIVPWQ-FVSQKTGATLKYVGLTKE-GVPDIEQ
ATHAL            GL--SNL----KPG-DEVILTVAEHHSCIVPWQ-IVSQKTGAVLKFVTLNED-EVPDINK
                       .        : :  :  ***. ::.**  :       :       .     :

HSALI            AADMITDDTALVNAVHISNTLGTVNPVGELADIAHDHGAYIFV--DGAQAAPTRAVDVQE
PHORI            AEKKIK-GAKLVAVQHVSNALGVIHEVEELGRMAKEEGAIFVV--DAAQSVGHMEVDVKK
ECOLI            LPTLFDEKTRLLAITHVSNVLGTENPLAEMITLAHQHGAKVLV--DGAQAVMHHPVDVQA
ECHRY            LPALLDERTRLVAVTQVSNVLGTVNPLAEIIRQAHACGAKVLV--DGAQAVMHQAVDVQA
NEURO            YQKLLNSHTRLVAFSQVSNALGTITPARQMVEMAHRVGARVLV--DGAQSVSHMRVDVQQ
LPNEU            FERKLNENTKMVAINYASNSLGTINPVKTMIKMAHEVGAKVLL--DGAQATAHLIVDVQD
BBACT            FKKKLNSRTKMVAFTASSNVLGTNTDMKLLTKLAHEVGAKVLV--DGAQIVSQLPVDVSD
ANABA            FKQLICEKTKLVSIVHVSNTLGCINPVAEIAEITHRYGAKFLV--DACQSVPHIPVDVQQ
MTUBE            L--YLDDRVKVVAFTHHSNVTGVLTPVSELVSRAHQSGALTVL--DACQSVPHQPVDLHE
BANTH            ARQTITPNTKIVSIMYVSNVLGTINPVKEIGAIAHENGAIMVV--DGAQSTPHMKVDVQD
PFALC            LISNMNINTKVISICHASNVIGNIQNIEKIIKKIKNVYPHIIIIDASQSFAHIKYDIKK
CMERO            FEKMLSERTKLVAIAYVGNVLGYVQDVRRITELAHQAGASVLV--DACQAVPHMPVDVQD
OSATI            LKGLLSNKTKIVVVHHVSNVLVSMLPIEDIVTWSNRIGAKVLV--DACQSVPHMPVDVQR
ATHAL            LRELISPKTKLVAVHHVSNVLASSLPIEEIVVWAHDVGAKVLV--DACQSVPHMVVDVQK
                     :   . ::     .*          :    :   .  .:  *..*       *:
```


HSALI IDA-----DFYAFSGHK-MLGPTGIGCLYGKRHLLA-E--MEPFLYGGDMIERVSY-EDA
PHORI LNA-----DFLAFSGHKGPMGPTGIGVLFISEEYFD-V--FDPPLIGGGTIEDVDL-CCY
ECOLI LDC-----DFYVFSGHK-LYGPTGIGILYVKEALLQ-E--MPPWEGGGSMIATVSLSEGT
ECHRY LDC-----DFYAFSGHK-LYGPSGIGVLYGKSELLQ-A--MPPWEGGGAMIREVSLTQGT
NEURO LDC-----DWFVFSGHK-VFGPTGIGVLYGKTELLN-D--TQPWQGGGNMIQDVTF-EKT
LPNEU LDC-----DFYAFSGHK-MYGPTGIGVLWGKEELLN-S--MTPYQGGGEMINSVSF-EAT
BBACT IDC-----DFFVFSAHK-LFGPFGFGAVYGKKEILD-Q--MPPYQGGGSMISKVTI-EKT
ANABA IDC-----DWLVASGHK-MCAPTGIGFLYGKLALLE-A--MPPFFGGGEMIAEVYL-DHS
MTUBE LGV-----DFAAFSGHK-MLGPNGIGVLYGRRELLA-Q--MPPFLTGGSMIETVTM-EGA
BANTH LNC-----DFYALSAHK-MCGPTGIGVLYGKKELLN-N--MEPIEFGGEMIDFVDL-QES
PFALC MKKNKSCPDILITSGHK-FCASLGTGFIFINKELSS-KYKFKPLLYGSNIITNVSK-YKS
CMERO IGC-----EWLVASGHK-MCGPTGIGILYGKEHVLRRE--MAPFLGGGEMIAEVFL-DHS
OSATI LGA-----DFLVASSHK-MCGPTGVGFLHGKFDLLS-S--MEPFLGGGEMIADVFQ-DKS
ATHAL LNA-----DFLVASSHK-MCGPTGIGFLYGKSDLLH-S--MPPFLGGGEMISDVFL-DHS
HSALI TW-NDPPWKFEAGTPVIAQGIALAEAVDYLQDI--GMDAIRAHEEALT--------EYAY
PHORI KL-TEPPERFEAGTPNIGGAIGLAAGIKYIEKI--GIEKIEKQERKLV--------KRTT
ECOLI TW-TKAPWRFEAGTPNTGGIIGLGAALEYVSAL--GLNNIAEYEQNLM--------HYAL
ECHRY TY-ADPPWRFEAGSPHVAGIIGLGAALDYVSAL--GVDAIQAHEGLLM--------RYAL
NEURO AY-HGAPARFEAGTGNIADAVGLGAAIDYVERI--GLENISRYEHELM--------TYAT
LPNEU EY-AAIPHKFEAGTPNIAGAIGLAAAIDYIWSL--DLDAIAEYETQLL--------NYAT
BBACT TF-NDVPTRFEAGTPHVEGAVGLHAALTFVENI--GLDKIHKYEMDLL--------NYAT
ANABA TY-AELPHKFEAGTPAIGEAIALGAAIDYLTNI--GMEQIHAYEAELT--------AYLY
MTUBE TY-APAPQRFEAGTPMTSQVVGLAAAARYLGAI--GMAAVEAHERELV--------AAAI
BANTH TW-KELPWKFEAGTPIIGNAIGLGAAIDFLEEI--GLDNIEKHEHELA--------QYAL
PFALC KFVTSLSELLETGTQNIPGILSMGISLEFFKKI--NWNYVYQYEMYLYDLFIYYMNKYMK
CMERO TF-ADLPHKFEAGTPSIGDAVALGAAIDYLEDGLGGMKRIHQFELQLA--------RYLY
OSATI TY-AEPPSRFEAGTPAIGEAIGLGTAIDYLSQI--GMQKIHEYENELA--------TYLY
ATHAL TY-AEPPSRFEAGTPAIGEAIALGAAVDYLSGI--GMPKIHEYEVEIG--------KYLY
HSALI DQLTMTDDVDVYGPPGD----------
PHORI EGLDELE-IPWYGPRNLD----------
ECOLI SQLESVPDLTLYGPQ----------
ECHRY ASLAEVPTLRLYGPV----------
NEURO GCLKSIAGLRLIGTAP----------
LPNEU KAIEAVKGYNIIGTAA----------
BBACT GRLLEIPDVKIYGTSQ----------
ANABA QQLEQIPQITLYGPKPDAN----------
MTUBE EGLSGIDGVRILGPTSMR----------
BANTH ERLSEVDGVTIYGPK----------
PFALC NHFVQLPNLNLSYKKENINYKSHMQTHPPVHKYNDEQNFTNDHNITQSKQTKSIHSQHDT
CMERO ESLEQFQEISIYGPPLDND----------
OSATI ESLIAVPNVRIYGPAPCQT----------
ATHAL EKLSSLPDVRIYGPRPSES----------
HSALI ----------DRGAVVSFNVDGIHAHDLSSILND-YGVAIRAGDHCTQPLHD-T-LG
PHORI ----------KHAGVVSFNVPPLHPHDVAAVLDE-HKIMVRSGHHCALPVMK-K-LG
ECOLI ----------NRLGVIAFNLGKHHAYDVGSFLDN-YGIAVRTGHHCAMPLMA-Y-YN
ECHRY ----------HRQGVIAFNLGRHHAFDVGSFLDQ-YGIAIRTGHHCAMPLMS-R-YG
NEURO ----------DKAGVLSFTLKDFSTEEVGTALNR-EGIAVRAGHHCAQPILR-R-FG
LPNEU ----------NKVPIISFVHGKIHAHDIGTILDS-EGIAIRSGHHCTMPLMD-F-YD
BBACT ----------NKGAILSFNLKGAHHSDIGQILDQ-EGVAVRAGHHCTQPLMA-R-LG
ANABA ----------GEGRAALATFTTTGVHANDLSTLLD-QEGVAIRSGHHCTQPLHR-H-LG
MTUBE ----------DRGSPVAFVVEGVHAHDVGQVLDD-GGVAVRVGHHCALPLHR-R-FG
BANTH ----------HRAGLVTFNIEDVHPHDVATVLDV-EGIAVRAGHHCAQPLMK-W-LK
PFALC FKIYTHDTRKYGLKKIGILPLWSNTFSSFDLVTFLDF-KNICIRAGHHCASLLHK-YYLK
CMERO ----------LGLERAALCAFNVRGVHPSDLATIIDL-DGIAIRAGHHCAQPLHRDA-LG
OSATI ----------VHRAPLCSFNVENVHPTDIAEILDLQHGVAIRSGHHCAQILHR-T-LG
ATHAL ----------VHRGALCSFNVEGLHPTDLATFLDQQHGVAIRSGHHCAQPLHR-Y-LG

```
HSALI           VPASARASFYLYNTRDEVDALVAAVDEARQIFAP-------
PHORI           INGTVRASFHVYNSVEEVEIFLGVLEDLVRKIRGAL-----
ECOLI           VPAMCRASLAMYNTHEEVDRLVTGLQRIHRLLG--------
ECHRY           VPSMCRASLALYSCQDEIDRLVAGLHRIHRLLGE-------
NEURO           VESTVRPSLAFYNTYTDIDRLAAAIDRIRGKKNAF------
LPNEU           VAATSRISMSFYNTFKEIDYCMEALQRVKEVFA--------
BBACT           VPGTVRASLSVYNNREDIDSMVKAVVKAREMLL--------
ANABA           LAATARASLSFYNTREDIDVFIKALKETIDFFAGIF-----
MTUBE           LAATARASFAVYNTADEVDRLVAGVRRSRHFFGRA------
BANTH           ASSTARASFYLYNTKEEIDTFVESLIKTKEYFTNVI-----
PFALC           VPDTSRISIYFYNTPQEIKYLAQQIASTSFMLNEMKNEK--
CMERO           VGGSARASVYVYNSSADIDRFIDALVDAVSVLGEKLTLRPA
OSATI           INASARASLHFYNTKEEVDVFVDALKDTIDFLTSEH-----
ATHAL           VNASARASLYFYNTKDDVDAFIVALADTVSFFNSFK-----
                     * *. .*.   ::.     :
```

This edited subset of sequences was aligned and bootstrapped trees constructed by parsimony; and by neighbour joining and maximum likelihood. A parsimony tree is presented (Fig. 2.4). *Plasmodium* spp. SufS was distantly placed from equivalents in bacteria, algae and vascular plants. This grouping had strong bootstrap support (99/100).

2.3.4.2. SufS-like sequences in prokaryotes.

The NCBI genome databases were searched using TBLASTN with EcSufS (GenBank AAL79956) as a template. This search revealed SufS to be present in a large number of bacterial and Archeal taxa. The presence of SufS in Archaea suggests the antiquity of the enzyme.

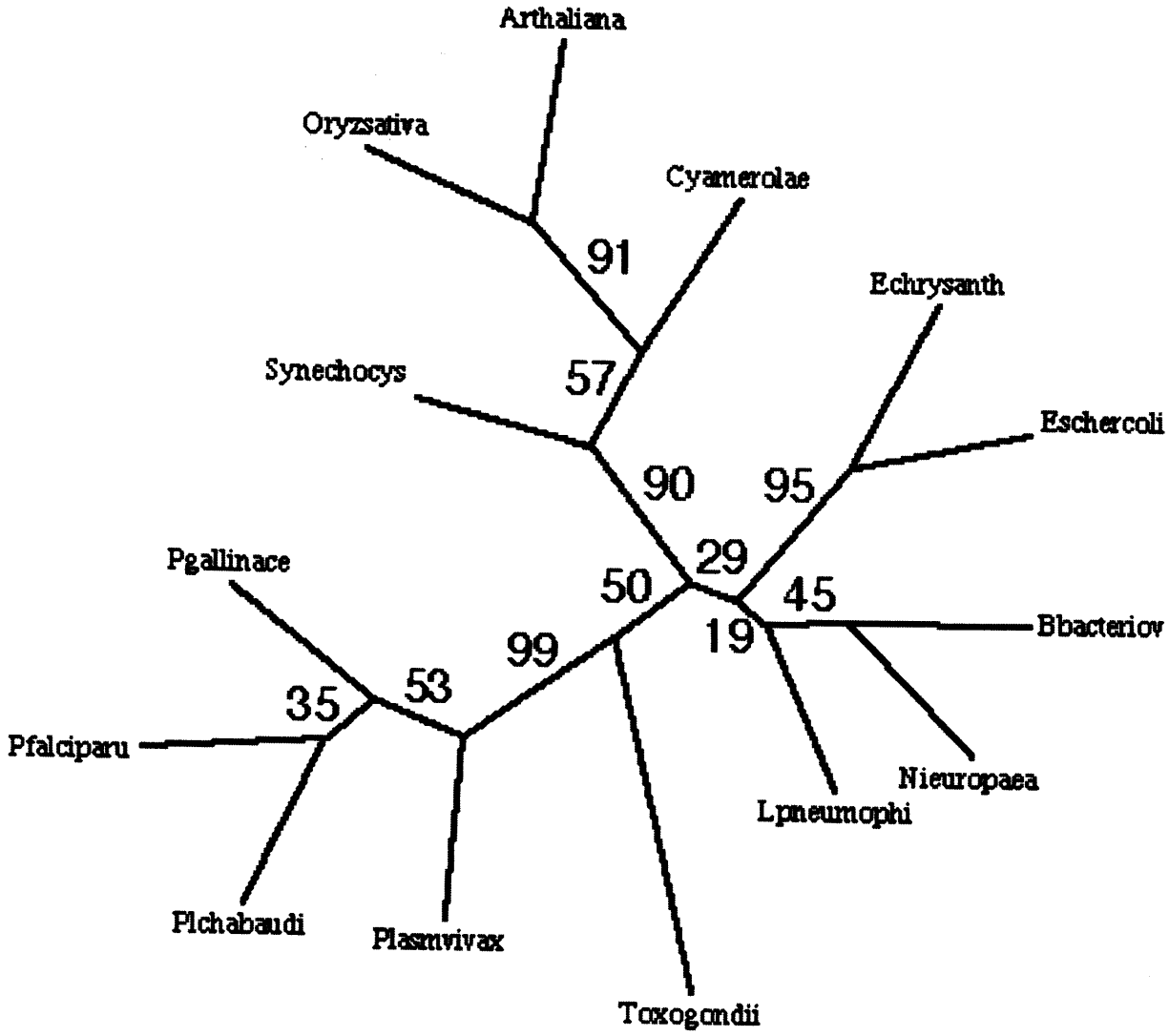
2.3.4.3. SufS sequences in algae and vascular plants.

Algal and plant genome databases were searched using TBLASTN with AtSufS (GenBank AAL79956) as a template. It is important to realise that these databases offer a biased taxonomic sample: for economic reasons, only the major crops, in particular the grasses (e.g., wheat), are covered. Nevertheless, it was found that protein sequences were conserved (identities > 95%) with AtSufS (data not shown). Additionally, SufS was identified in the red alga, *Cyanidioschyzon merolae*. Notable vascular plant examples include *Lycopersicon esculentum* (Solanaceae), *Gossypium barbadense* (Malvaceae), *Citrus sinensis* (Rutaceae) and *Pinus taeda* (Pinaceae). In conclusion, SufS is probably found in all photosynthetic eukaryotes.

FIGURE 2.4.

PHYLOGENETIC TREE OF SUFS PROTEIN SEQUENCES.

A phylogenetic tree was constructed from SufS protein sequences by parsimony. Bootstrap values are given (100 pseudoreplicates). Abbreviations are as follows. Echrysanth: *E. chrysanthemi*; Eschercoli: *E. coli*; Bbacteriov: *B. bacteriovorus*; Nieuropaea: *N. europaea*; Lpneumophi: *L. pneumophila*; Toxogondii: *T. gondii*; Plasmvivax: *P. vivax*; Plchabaudi: *P. chabaudi*; Pfalciparu: *P. falciparum*; Pgallinace: *P. gallinaceum*; Synechocys: *Synechocystis* sp.; Oryzsativa: *O. sativa*; Arthaliana: *A. thaliana*; Cyamerolae: *C. merolae*.


Arthaliana
Oryzsativa
Cyamerolae
91
Echrysanth
Synechocys
57
Eschercoli
90
95
29
Pgallinace
50
45
Bbacteriov
19
99
53
35
Pfalciparu
Nieuropaea
Lpneumophi
Plchabaudi
Plasmvivax
Toxogondii

2.3.4.4. SufS-like sequences in the Apicomplexa.

TBLASTN searching of genome databases with EcSufS revealed complete or partial SufS sequences in several plastid-bearing Apicomplexans. For example, *P. falciparum*, *P. vivax*, *P. knowlesi*, *P. berghei*, *P. chabaudi chabaudi*, *P. yoelii* and *P. gallinaceum.*

P. yoelii SufS (PySufS) sequence has been assembled *in silico* and checked by RT-PCR (Ellis and Clough, NIMR, unpublished observation). Additionally, *P. chabaudi chabaudi* (PC302264.00.0) and *P. berghei* (PB000547.03.0) SufS sequences were assembled by Hall, *et al.*[89] I manually assembled the sequence of *P. vivax* SufS (PvSufS) and PfSufS to supplement those assembled automatically in PlasmoDB, namely, PfSufS (PF07_0068), PySufS (PY06105) and PvSufS (Pv_4040.phat_55). However, these bioinformatic assemblies must be viewed with extreme caution as errors in intron boundaries and start sites have been noted in several instances (Dr. B. Clough, pers. comm.)

With regard to other Apicomplexan organisms, partial amino acid sequence of *T. gondii* SufS was assembled *in silico* from genomic read TGG_10447 (Sato, NIMR, unpublished observation). Although often classified outside the Apicomplexa, the oyster pathogen, *Perkinsus marinus*, unambiguously encodes SufS. Additionally, SufS hits were evident in the *Babesia bovis, Eimeria tenella* and *Theileria annulata* databases; but large gaps prevented reliable assembly of sequences. Notably, SufS is absent from Apicomplexans lacking a plastid, namely, *C. hominus* and *C. parvum.*

In an effort to quantify the high level of conservation between the Apicomplexan SufS sequences, 45 amino acids C-terminal of the SXGK motif were selected for estimation

of the ratio of rates of synonymous to non-synonymous mutations (dS/dN). This parameter gives an indication of the force of selection upon protein-coding DNA.[229] DNA encoding the region was aligned using the T-COFFEE algorithm and refined by eye; dS/dN ratios were calculated as described (section 2.2.1.1). Unsurprisingly, dS/dN ratios were low (< 0.2), hinting at the action of purifying selection. The most convergent pair were those of *P. falciparum* and *P. gallinaceum* (0.0029); while the most divergent pair were those of *P. knowlesi* and *P. vivax* (0.1827).

2.3.5. Description of PfSufS.

A detailed description of PfSufS will now be given. Messenger RNA predicted to encode the enzyme is 1641 nucleotides (nt) long and would translate into a protein of 546 amino acids (*ca.* 65 kDa). The gene is composed of two exons (291 nt and 1350 nt) divided by one intron (283 nt). The AU content of the mRNA is 76 %, typical of *P. falciparum* (the AT bias has implications for cloning and heterologous expression – see chapter 5). Exon one encodes a leader sequence and a short stretch of the mature protein. The leader sequence is predicted to target the protein to the plastid (chapter 3). Exon 2 encodes the aminotransferase type V domain containing the PLP-binding and active site motifs (more detail is given in chapter 4).

Inspection of the alignment of PfSufS with EcSufS (Fig. 2.3) indicates regions of significant difference. Chief among these were two insertions: a penta-isoleucine (5I) tract (I255-I259) and a 51-amino acid tract. With regard to the 5I tract, it is worth noting that similar insertions enriched for non-polar hydrophobic amino acids (IVL) were present in all other Apicomplexan sequences examined. Furthermore, low dS/dN ratios suggest that this feature is under functional constraint (i.e., purifying selection).[230]

To investigate further the properties of the 51-amino acid tract, a PSI-BLASTP search of the NCBI database was conducted using the *P. falciparum* tract as a probe. Weak hits to rodent prion proteins (PrP) were produced. Hence, the tract was designated as the "PrP-like" loop. Alignment of the PrP-like loop of PfSufS with authentic PrP sequences is presented (Fig. 2.5).

Although insertions equivalent to the PrP-like loop were present in other Apicomplexan SufS sequences, amino acid composition and length was not conserved (data not shown). Indeed, corresponding nucleotide sequences could not be aligned; hence dS/dN ratios could not be calculated (data not shown). However, the edges of the loop showed greater sequence conservation than the central region, as has been reported for loops in *P. falciparum* and *P. berghei* γ-glutamylcysteine synthetases.[231] It is worth noting that insertions are common in *P. falciparum* proteins.[232] Taken together, these data imply a lack of functional constraint. It is tempting to speculate on a chance insertion event after the divergence of Apicomplexa from the bacterial, algal and plant lineages.

2.3.6. Physical structure of PfSufS.

A hydropathy plot of PfSufS (Fig. 2.6) revealed a hydrophobic N-terminal signal sequence that may be involved in targeting the protein to the plastid organelle (chapter 3). PfSufS also appeared to lack transmembrane regions. In collaboration with Dr. J. W. Saldanha, a molecular homology model of PfSufS was developed based on the X-ray crystal structure of EcSufS (PDB 1JF9). It should be noted that the first 60 amino acids of PfSufS and the PrP-like loop were excluded from the model.

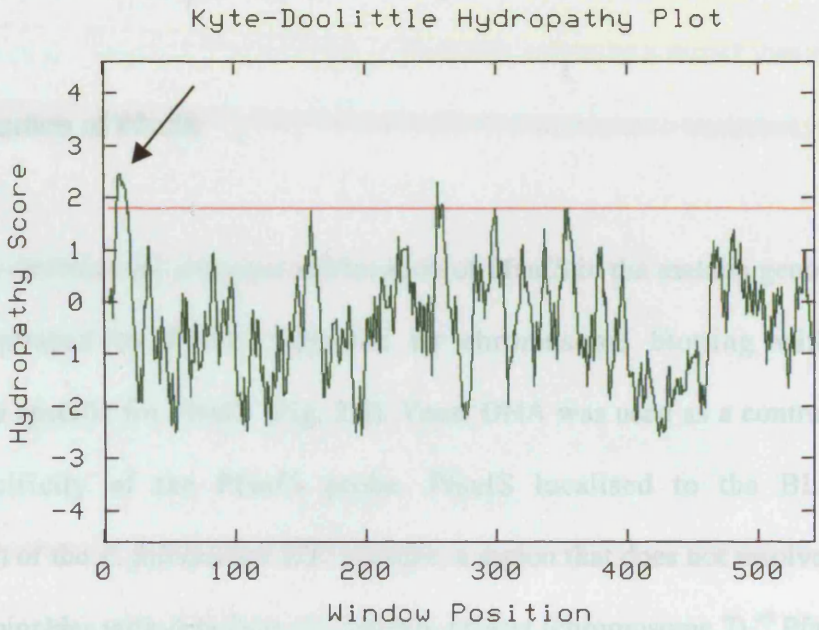
FIGURE 2.5.

ALIGNMENT OF PFSUFS PRP-LIKE LOOP WITH AUTHENTIC PRION PROTEINS.

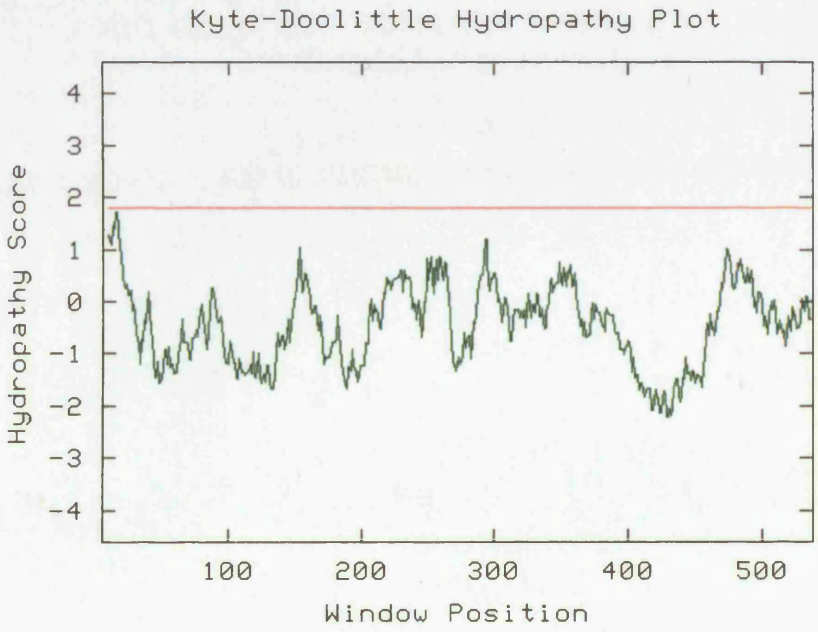
Sequences were retrieved from the NCBI database using the PSI-BLASTP algorithm with a portion of the *P. falciparum* PrP-like loop as probe. Weak hits (E = 24-27) were made with PrP in three hamster taxa: *Cricetulus migratorius* (GenBank Q60468); *C. griseus* (GenBank Q60506); and *Mesocricetus auratus* (GenBank P04273). These sequences were aligned with the PrP-like loop of PfSufS. Also included in the alignment were PrP sequences from *Mus musculus* (GenBank P04925) and *Homo sapiens* (GenBank P04156). Abbreviations are as follows. PRIO_CRIMI: *C. migratorius*; PRIO_CRIGR: *C. griseus*; PRIO_MESAU: *M. auratus*; PRIO_MOUSE: *M. musculus*; PRIO_HUMAN: *H. sapiens*; and SUFS_PFALC: *P. falciparum.*

```
CLUSTAL FORMAT for T-COFFEE Version_1.41, CPU=20.42 sec, SCORE=49, Nseq=6, Len=255

PRIO_CRIMI      MANLSYWLLALFVATWTDVGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGGTWGQP
PRIO_CRIGR      MANLSYWLLALFVATWTDVGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGGTWGQP
PRIO_MESAU      MANLSYWLLALFVAMWTDVGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGGTWGQP
PRIO_MOUSE      MANLGYWLLALFVTMWTDVGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGG-TWGQP
PRIO_HUMAN      MANLGCWMLVLFVATWSDLGLCKKRPKPGGWNTGGSRYPGQGSPGGNRYPPQGGGGWGQP
SUFS_PFALC      ------------------------------------------------------------


PRIO_CRIMI      HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHNQWNKPNKPKTSMKHMAGAAAAGA
PRIO_CRIGR      HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHNQWNKPSKPKTNMKHVAGAAAAGA
PRIO_MESAU      HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHNQWNKPSKPKTNMKHMAGAAAAGA
PRIO_MOUSE      HGGGWGQPHGGSWGQPHGGSWGQPHGGGWGQGGGTHNQWNKPSKPKTNLKHVAGAAAAGA
PRIO_HUMAN      HGGGWGQPHGGGWGQPHGGGWGQPHGGGWGQGGGTHSQWNKPSKPKTNMKHMAGAAAAGA
SUFS_PFALC      ------------------------------------------------------------


PRIO_CRIMI      VVGGLGGYMLGSAMSRPMLHFGNDWEDRYYRENMNRYPNQVYYRPVDQYNNQNNFVHDCV
PRIO_CRIGR      VVGGLGGYMLGSAMSRPMLHFGNDWEDRYYRENMNRYPNQVYYRPVDQYNNQNNFVHDCV
PRIO_MESAU      VVGGLGGYMLGSAMSRPMMHFGNDWEDRYYRENMNRYPNQVYYRPVDQYNNQNNFVHDCV
PRIO_MOUSE      VVGGLGGYMLGSAMSRPMIHFGNDWEDRYYRENMYRYPNQVYYRPVDQYSNQNNFVHDCV
PRIO_HUMAN      VVGGLGGYMLGSAMSRPIIHFGSDYEDRYYRENMHRYPNQVYYRPMDEYSNQNNFVHDCV
SUFS_PFALC      -----------------------------KKENINYKSHMQTHPPVHKYNDEQNFTNDH-
                                              :**:   .:   : *:.:*.:::**.:*


PRIO_CRIMI      NITIKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYQKESQAYYDGRRSSA-VLFSSP
PRIO_CRIGR      NITIKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYQKESQAYYDGRRSSA-VLFSSP
PRIO_MESAU      NITIKQHTVTTTTKGENFTETDIKIMERVVEQMCTTQYQKESQAYYDGRRSSA-VLFSSP
PRIO_MOUSE      NITIKQHTVTTTTKGENFTETDVKMMERVVEQMCVTQYQKESQAYYDGRRSSSTVLFSSP
PRIO_HUMAN      NITIKQHTVTTTTKGENFTETDVKMMERVVEQMCITQYERESQAYYQ-RGSSM-VLFSSP
SUFS_PFALC      NITQSKQTKSIHSQHDTFKIYTHDTRKYGLKKI---------------------------
                *** .::* :  :: :.*.    .  :  ::::

PRIO_CRIMI      PVILLISFLIFLIVG
PRIO_CRIGR      PVILLISFLIFLIVG
PRIO_MESAU      PVILLISFLIFLMVG
PRIO_MOUSE      PVILLISFLIFLIVG
PRIO_HUMAN      PVILLISFLIFLIVG
SUFS_PFALC      ---------------
```

Window size = 9



Window size = 19



The first 60 amino acids approximately correspond to the plastid-targeting leader; while the PrP-like loop was said to be unstructured and therefore could not be modelled (Dr. J. W. Saldanha, pers. comm.). The molecule (Fig. 2.7) contains 276 hydrogen bonds; 20 helices; 21 strands; 43 turns; but lacks disulphide bridges.

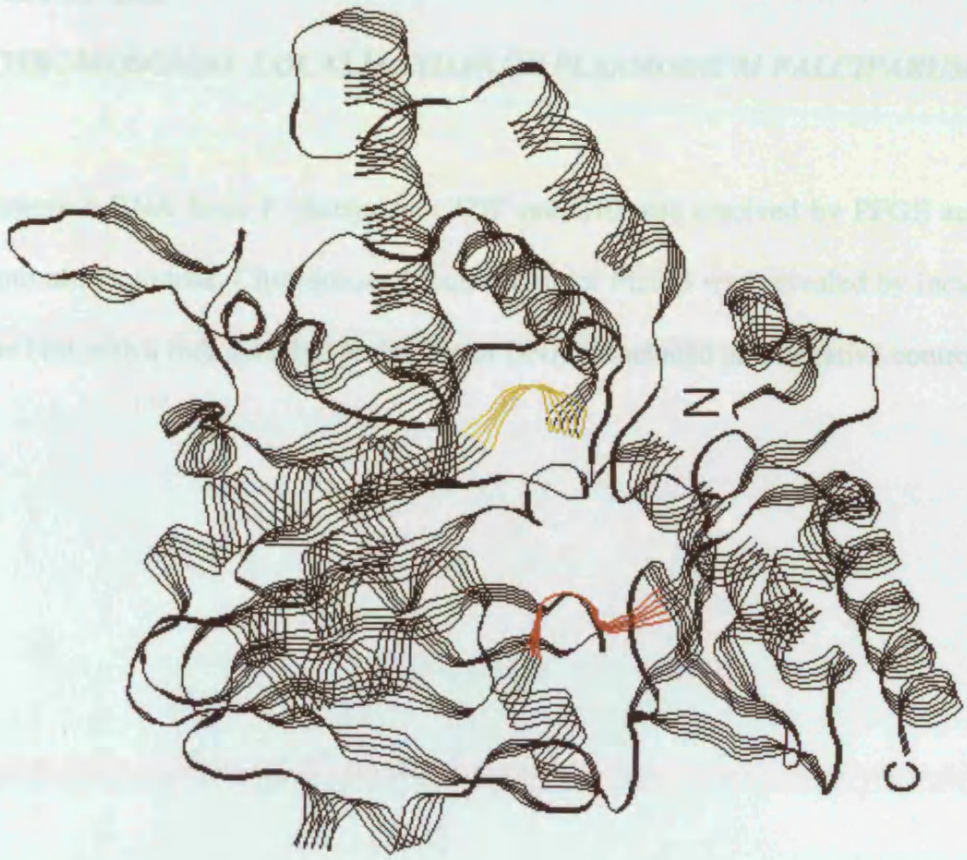
2.3.7. Genomic location of PfsufS

It was important to establish the presence and location of PfsufS in the malaria genome. This was accomplished by PFGE, followed by chromosome blotting with a radiolabelled probe specific for PfsufS (Fig. 2.8). Yeast DNA was used as a control to confirm the specificity of the PfsufS probe. PfsufS localised to the BLOB (chromosomes 6-9) of the *P. falciparum* 3D7 genome, a region that does not resolve on PFGE gels. This coincides with data from the genome project (chromosome 7).[45] PfsufS is also encoded by the *P. falciparum* C10 strain. However, in C10, the genomic location cannot be precisely defined because C10 chromosomes are heavily fragmented (Dr. D. H. Williamson, NIMR, pers. comm.).

FIGURE 2.7.

MOLECULAR MODEL OF *PLASMODIUM FALCIPARUM* SUFS.

Stereoscopic view into the active site cleft of the enzyme. Lysine-binding (yellow) and catalytic (red) motifs are marked. The enzyme is probably active as a dimer, but only the monomers are shown. Plastid-targeting leader and PrP-like loop are excluded from the model.


Z
C

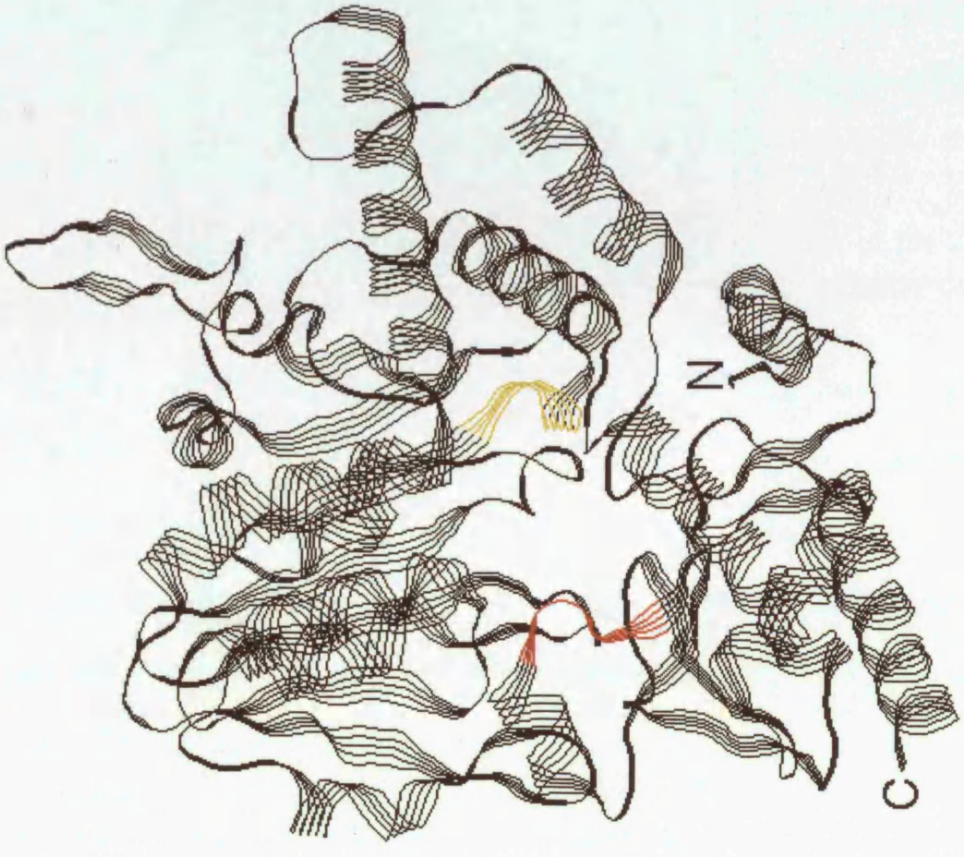
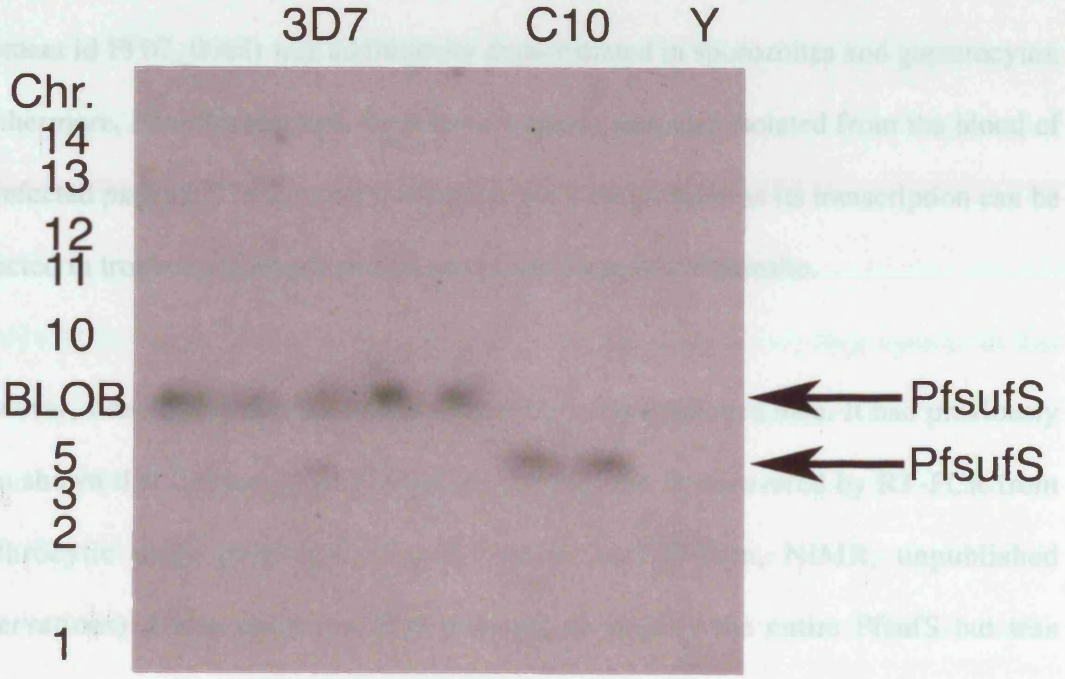

Z
C

FIGURE 2.8.

CHROMOSOMAL LOCALISATION OF *PLASMODIUM FALCIPARUM* SUFS.

Genomic DNA from *P. falciparum* 3D7 and C10 was resolved by PFGE and blotted onto nitrocellulose. Chromosomal localisation of PfsufS was revealed by incubation of the blot with a radiolabelled probe. Yeast DNA is included as a negative control (Y).


3D7
C10
Y
Chr.
14
13
12
11
10
BLOB
5
3
2
1
PfsufS
PfsufS

2.3.8. Gene expression studies of PfsufS.

2.3.8.1. Transcription of PfsufS.

Microarray data for *P. falciparum* are available. Bozdech, *et al.*[18] detected transcription of PfsufS (element id oPFBLOB0139) throughout the erythrocytic stage, but with an upregulation in trophozoites (24 h); this result has been reproduced in 3D7 and Dd2 strains (DeRisi, UCSF, USA, unpublished observations*). In a study of *P. falciparum* HB3 developmental stages conducted by Le Roch, *et al.*,[83] upregulation of *PfsufS* (element id PF07_0068) was additionally demonstrated in sporozoites and gametocytes. Furthermore, *PfsufS* transcripts have been found in parasites isolated from the blood of an infected patient.[234] In summary, PfsufS is not a pseudogene as its transcription can be detected in trophozoite stages as well as sexual stages of the parasite.

However, I decided it was necessary to confirm these published data. It had previously been shown that full-length *PfsufS* transcript could not be recovered by RT-PCR from erythrocytic stage parasites (Clough, Slatter and Wilson, NIMR, unpublished observations). I also made repeated attempts to amplify the entire PfsufS but was unsuccessful (data not shown). This suggested that *PfsufS* might not be expressed in erythrocytic stage parasites manipulated in our hands. To test this hypothesis, PfsufS sequence was broken in half and two sets of primers designed. Total RNA was recovered from mixed erythrocytic stage cultures of *P. falciparum* 3D7. RNA was prepared fresh; a typical stock concentration was 1.6 $\mu g.\mu L^{-1}$. Crucially, the first set of primers spanned the intron boundary. Using these primers, RT-PCR was successful,

*http://plasmodb.org/restricted/gsdata/viewGlassSlideData.php?genesOnThisPage=1&geneID=PF07_0068

producing products of predicted size (Fig. 2.9). This demonstrates that *PfsufS* transcripts are present in erythrocytic stage parasites.

2.3.8.2. Models of regulation of SUF genes.

Two mechanisms of SUF regulation may exist (chapter 1). In *E. coli*, correlation was noted between oxidative stress, iron limitation and SUF expression (chapter 1). Thus, *E. coli* SUF genes are thought to be involved in the iron starvation or oxidative stress responses. Recall that oxidative stress and iron status are intimately related.[235] Indeed, Outten, *et al.*[178] consider the oxidative stress response of SUF genes to be ultimately due to iron limitation (chapter 1). However, in this study, the two conditions were investigated separately.

In eukaryotes and some bacteria, a different regulatory mechanism may operate to that found in *E. coli*. Specifically, the SUF system may fulfil a solely constitutive function. For example, in *P. falciparum*, the SUF system is predicted to supply [Fe-S] clusters exclusively to proteins located in the plastid organelle.[236] However, upregulation of specific components of the SUF system could still occur in response to [Fe-S] cluster damage.

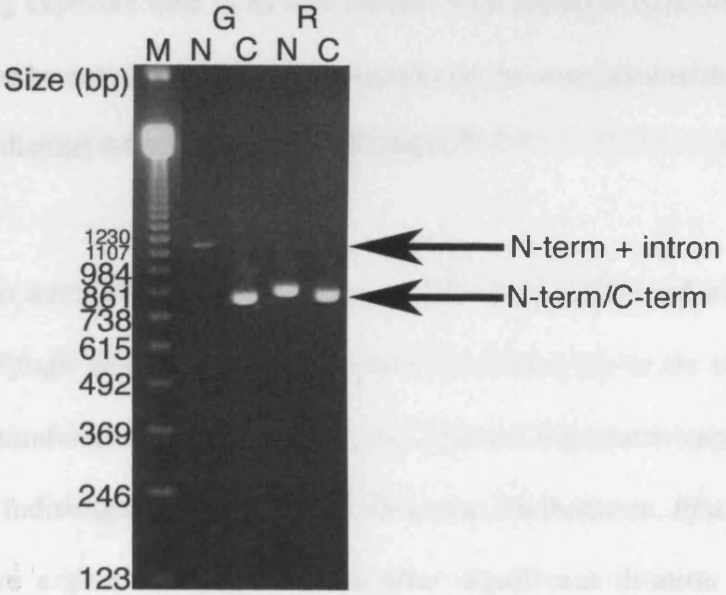
2.3.8.3. Design of assay.

In an effort to distinguish between these two possible models (regulated *vs.* constitutive) in *P. falciparum*, it was decided to investigate *PfsufS* mRNA levels in response to both oxidative stress and iron limitation.

FIGURE 2.9.

TRANSCRIPTION OF *PLASMODIUM FALCIPARUM* SUFS.

PfsufS[M] was recovered by PCR from genomic DNA (G) and cDNA (R) isolated from *P. falciparum* 3D7. Prior to RT-PCR, genomic DNA contamination was removed from RNA preparations by treatment with TURBO DNAse. Other abbreviations are as follows. bp: base pair; M: 123 bp ladder (Sigma); N: 5′ end of PfsufS[M]; C: 3′ end of PfsufS[M].


G
R
M N C N C
Size (bp)
1230
1107
984
861
738
615
492
369
246
123
N-term + intron
N-term/C-term

At the commencement of this work, this was the first time such transcriptional responses had been measured in *P. falciparum.* Hence, the initial step was to establish a simple assay. This process is now described.

Genes can upregulate rapidly, often within minutes. Five min probably approaches the minimum time for upregulation of a eukaryotic gene, e.g., leu-2 of *Neurospora crassa.*[237] In many gene expression studies, organisms are exposed to oxidative stress for short periods < 10 min.[238] However, given the necessity for the stressor agent to diffuse through the multiple membranes surrounding malaria parasites in the red blood cell, a long exposure time (3 h) was chosen. With regard to oxidative stress in *E. coli*, it should also be noted that SUF upregulation may be associated with prolonged oxidative attack, as distinct from short periods of stress.[146]

In order to normalise loading of RNA on RNAP gels, *PfactinI* was used as a control. Initially, *PfsufS* and *PfactinI* were probed simultaneously in the same tube. However, this was uninformative, because degraded *PfactinI* fragments were of a similar size to (and thus indistinguishable from) PfsufS probe. Furthermore, *PfactinI* transcripts were in massive excess to *PfsufS*. Even after significant dilution of PfactinI probe, autoradiograms rapidly saturated. Therefore, it was decided to probe the two genes in separate tubes. Indeed, RNA level could be accurately estimated merely by inspection of 28S and 18S ribosomal bands on agarose gel stained with ethidium bromide (Fig. 2.10A).

2.3.8.4. Oxidative stress response.

Parasites were placed under oxidative stress by culturing with H_2O_2. Toxicity of H_2O_2 to *P. falciparum* is marked. For example, exposure of parasite cultures to high concentrations of H_2O_2 (> 100000 ppm) rapidly lead to cellular lysis. Frothing was also observed, indicating host catalase activity. Thus, low levels (10-1000 ppm) of oxidant were used in all assays. In spite of a weak signal, it appeared that exposure of parasites to H_2O_2 for 3 h had no influence on *PfsufS* transcription (Fig. 2.10A).

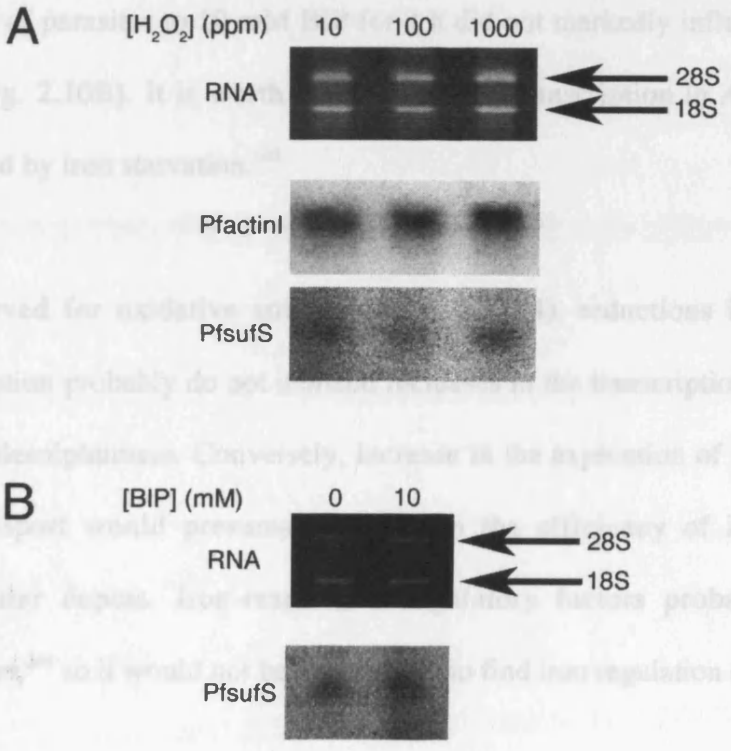
P. falciparum is known to be under considerable oxidative stress, notably, in the erythrocytic, ookinete and sporozoite stages; thus, the parasite has a highly developed antioxidant system.[239] However, *PfsufS* does not appear to be part of this system.

Returning to first principles, oxidative stress must have a predominantly demetalling influence on [Fe-S] clusters (chapter 1). Indeed, transcription of *sufS* was unchanged under oxidative conditions in *Synechocystis* sp.[240] Close examination of the *E. coli* microarray oxidative stress data[175] indicates that *EcsufA* upregulated significantly (21 ×), about an order of magnitude greater than the trivial increase in EcsufS transcription over basal levels (3.5 ×). The extent of *EcsufBCDE* upregulation was intermediate (8-16 ×). Therefore, in line with the hypothesis that oxidative stress preferentially damages the iron component of [Fe-S] clusters, it is tempting to speculate that SufA in particular operates in remetalling in *E. coli.*

FIGURE 2.10.

TRANSCRIPTIONAL RESPONSE OF *PLASMODIUM FALCIPARUM* SUFS TO OXIDATIVE STRESS AND IRON STARVATION.

Erythrocytic stages of *P. falciparum* were exposed to oxidative stress by treatment with 10, 100 and 1000 parts per million (ppm) H_2O_2 (A); or iron starvation by treatment with 0 or 10 mM of BIP (B). Total RNA was extracted (RNA) and probed for *PfactinI* and *PfsufS* by RNAse protection assay. Ribosomal bands (28S and 18S) are marked with arrows.


A
[H2O2] (ppm)
10
100
1000
RNA
28S
18S
PfactinI
PfsufS
B
[BIP] (mM)
0
10
RNA
28S
18S
PfsufS

2.3.8.5. Iron limitation response.

To investigate the possibility that PfsufS transcription responded to iron limitation, cultures were treated with BIP, a lipophilic compound that chelates Fe^{2+}, even in the cytosol of *P. falciparum* red blood cell stages.[241] BIP is toxic to *P. falciparum* (IC_{50}-24 h of 27.9-37.9 μM[242]) and has been proposed as an adjunct to conventional antimalarial chemotherapy.[243] The mechanism of toxicity is unknown. In my hands, BIP had a significant negative effect on the growth of *P. falciparum* (data not shown). However, exposure of parasites to 10 mM BIP for 3 h did not markedly influence *PfsufS* mRNA level. (Fig. 2.10B). It is worth noting that *sufS* transcription in *A. thaliana* was also unaffected by iron starvation.[185]

As observed for oxidative stress (section 2.3.8.4), reductions in intracellular iron concentration probably do not demand increases in the transcription of genes encoding cysteine desulphurases. Conversely, increase in the expression of proteins involved in iron transport would presumably heighten the efficiency of iron recovery from intracellular depots. Iron-responsive regulatory factors probably operate in *P. falciparum*,[244] so it would not be implausible to find iron regulation of SUF proteins.

2.3.8.6. Conclusion and future work.

With regard to the oxidative stress (section 2.3.8.4) and iron limitation (section 2.3.8.5) data, I found no evidence that PfsufS transcription was influenced by either of these environmental perturbations. This result is consistent with findings in bacteria and vascular plants (sections 2.3.8.4-5). However, specific attack on the sulphur component of [Fe-S] clusters, e.g., by NO (g), would presumably lead to upregulation of *sufS*

(chapter 1). Additionally, it seems possible that oxidative stress and iron limitation could regulate *PfsufA*. In future work, it would be fruitful to investigate these ideas using the assay reported here. Nevertheless, until a greater understanding of transcriptional control in *P. falciparum* is developed, the value of such simple experiments can be challenged.

CHAPTER 3.

EXPERIMENTS ON THE INTRACELLULAR LOCALISATION OF *PLASMODIUM FALCIPARUM* SUFS.

P. falciparum SufS is predicted to target the plastid. This chapter begins with a summary of current ideas on the mechanisms of plastid localisation in Apicomplexan cells. Prediction of the plastid localisation of PfSufS is described in detail. Investigations of this prediction, by transfection and immunolocalisation, are reported.

3.1. INTRODUCTION.

3.1.1. Protein traffic to the plastid in malaria.

3.1.1.1. The need for protein import to the plastid.

As soon as the minimal size of the plastid genome of *P. falciparum* (35 kb)[53] was known, it became apparent that the majority of plastid functions must be performed by heterochthonous proteins (c.f., plastids of vascular plants and algae): genes encoded in the nucleus are translated on cytosolic ribosomes; gene products are trafficked to the plastid where they act. The most recent bioinformatic estimate for the number of such genes in *P. falciparum* is 545 (encoding ~ 5 % of total proteome).[80] Understanding the function of the plastid requires the identification of proteins that traffic to it; hence, considerable effort has been expended in dissecting the phenomenon.[60,76,257]

3.1.1.2. Complexity of trafficking imposed by multiple membranes.

The plastid is bounded by multiple membranes. In *P. falciparum*, the number is either three or four, depending upon interpretation. For example, Hopkins, *et al.*[59] observed three bounding membranes by electron microscopy (EM). Conversely, four membranes were observed by EM of other Apicomplexan cells, e.g., *T. gondii*[54] and the reptile parasite *Garnia gonadoti* (Haemosporina: Garniidae).[258] However, it has been pointed out that the fourth membrane of the *T. gondii* plastid could be an extension of the endoplasmic reticulum (ER).[259] Survey of the Apicomplexan literature by Waller and McFadden[60] apparently revealed a four-membrane plastid in *Hepatozoon domerguei* (Coccidia: Adeleidea),[260] *Coelotropha durchoni* (Coccidia: Grelliidae)[261] and *Selenidium*

pendula (Sporozoa: Archigregarinida).[262] Resolution of the number of membranes is important, as it has significant implications for the mode of plastid trafficking.

The inner two membranes of the plastid are believed to be derived from the primary endosymbiont (chapter 1) and hence may be termed primary inner and outer membranes. These membranes have, in turn, been encircled by one (or two) membrane(s) from secondary endosymbiosis (secondary inner and outer membranes). The secondary inner membrane may correspond to the cell membrane of the secondary endosymbiont, while a secondary outer membrane could be a vestige of the endocytic vesicle membrane. The compartment between the primary and secondary membrane layers is termed the periplastid space.

3.1.1.3. Mechanism of intracellular trafficking to the plastid.

One experimental paradigm for plastid trafficking is acyl carrier protein (ACP) in the erythrocytic stages of *P. falciparum* (PfACP; PlasmoDB PFB0385w) and in the tachyzoite stages of *T. gondii* (TgACP; GenBank AAC63956). Experiments conducted with ACP indicated that nuclear-encoded proteins addressed to the plastid can be recognised by an N-terminal bipartite leader sequence (ACP[L]) composed of a signal peptide and a plastid transit peptide.[76] This mode of targeting seems to apply generally to other parasite plastid proteins.

Fluorescent reporters (e.g., GFP) can be addressed to the plastid *in vivo* by genetic fusion to an N-terminal leader sequence. By this method, it was shown that specific ablation of the signal peptide led to accumulation of fluorescence in the cytosol; conversely, ablation of the plastid transit peptide led to accumulation of fluorescence in

the parasitophorous vacuole.[79] Hence, it was assumed that plastid-targeted proteins were first directed to the secretory pathway, presumably via the Golgi apparatus (at least in *T. gondii*, where a Golgi is clearly observed), before diversion to the plastid.

However, it has been claimed that proteins also pass through the parasitophorous vacuole *en route* to the plastid.[263] Additionally, inhibition of Golgi trafficking by treatment of *T. gondii* with brefeldin-A, or incubation at low temperature (15 °C), did not abolish plastid import[264] Furthermore, appending a *T. gondii* ER retrieval (HDEL) signal to the C-terminal of the GFP constructs did not influence plastid targeting.[264] These data argue against involvement of the classical secretory pathway.

Nevertheless, as a generalisation, the signal peptide is thought to be cleaved in the ER. Vesicles[265] may convey the partially cleaved protein, possibly bound to the chaperone DnaK (section 3.1.2.3), and dock to components of the outermost membrane of the plastid. This could lead to insertion of the protein into the space underneath the outermost membrane. If the plastid is enveloped by four membranes (section 3.1.1.2), the protein would face the additional challenge of passing through a second membrane to reach the periplastid space. It is possible that this could be achieved via interaction with a duplicated Toc complex,[76] but the actual mechanism is unclear.

DeRocher, *et al.*[264] have proposed the existence of a pre-plastid compartment in which interaction between the C-terminal HDEL signal of the plastid-targeted protein and an ER retrieval receptor (e.g., Erd2) could occur (leading to re-entry of the protein into the secretory pathway). Presumably, for successful import to the plastid, this retrograde process must be out-competed by interaction between the plastid transit peptide and

receptors on the plastid membrane. One speculation is that this pre-plastid compartment corresponds to the periplastid space.

Following entry into the periplastid space, the import pathway is believed to mimic that of the chloroplast. Specifically, a Tic/Toc translocon complex[266] may mediate import through the two primary membranes, possibly energised by the autochthonous ATPase, PfClpC (GenBank S72278).[236] However, it is worth noting that only a putative Tic22 orthologue has been shown to target the plastid in *P. falciparum* (Tonkin and McFadden, University of Melbourne, Australia, unpublished observation) and that no other Tics/Tocs have been identified. Cleavage of the transit peptide is probably achieved by endopeptidases in the plastid lumen (e.g., PfSPP; GenBank AF453250), releasing mature protein that can adopt a functional conformation, perhaps by the action of imported chaperones, e.g., Cpn60.[267]

3.1.1.4. Kinetics of trafficking.

By pulse-chase experiments with PfACP[L]-GFP, van Dooren, *et al.*[77] estimated the minimum time from protein synthesis, to cleavage and import into the plastid to be 30-40 min. This is similar to the import time for the small subunit of ribulose-15-bisphosphate carboxylase/oxygenase to the secondary plastid of *Euglena gracilis* (*ca.* 40 min).[268] However, it is considerably longer than for import of the L18 ribosomal protein to the primary plastid of the green alga *Chlamydomonas reinhardtii* (< 10 min).[269] The rate-limiting step for import of proteins to the Apicomplexan plastid is thus assumed to be the cleavage of the plastid transit peptide.

These kinetics give a characteristic delay signature on Western blots. When *P. falciparum* lysate is probed with antibody to plastid-targeted protein, a low abundance high molecular weight band (partially cleaved protein) and a high abundance low molecular weight band (mature protein) are seen.[77,97]

3.1.1.5. Comparison with trafficking to other secondary plastids.

Plastid trafficking in Apicomplexa is often modelled on the mechanisms in vascular plants. However, although data are sparse, it is probably more useful to look for clues in organisms containing secondary plastids. Important examples of cells hosting plastids bound by three membranes include the alga *Euglena*, while chromists harbour a plastid bounded by four membranes.[265] However, based on a molecular phylogenetic study of duplicated GAPDH genes,[62] dinoflagellates are considered the most closely related taxon to the Apicomplexa, with triple-membrane plastids. Nassoury, *et al.*[270] studied plastid trafficking in the dinoflagellate *Gonyaulax polyedra* (Dinophyceae: Gonyaulacaceae) and revealed that a bipartite leader was required for plastid targeting, and that traffic passed through the Golgi apparatus (not a prominent organelle in *P. falciparum* erythrocytic stages).

3.1.2. Minimum determinants for plastid targeting in Apicomplexa.

3.1.2.1. The signal peptide.

The signal peptide portion of the leader sequence of plastid-targeted proteins is of the classical hydrophobic type and hence readily detected using standard algorithms, e.g.,

SignalP. The signal peptide is probably co-translationally cleaved as the protein is drawn into the parasite ER, revealing the plastid transit peptide.

3.1.2.2. The plastid transit peptide.

In common with chloroplast transit peptides, the plastid transit peptide portion of the leader exhibits a net positive charge. However, unlike equivalents in vascular plants, it is enriched for asparagine and lysine and depleted in acidic residues.[79] Mutation experiments with PfACP[L] indicated that the net positive charge was absolutely required for plastid-targeting.[79] Besides these properties, there appears to be considerable functional redundancy in primary sequence.[271] Thus, mutations of the plastid transit peptide might be expected to be under limited functional constraint and therefore subject to faster rates of base substitution (as compared to sequence coding for the mature protein). Indeed, it has been suggested that the major constraining influence on the amino acid composition of the plastid transit peptide is the codon bias of the DNA.[272] All this adds up to a lack of association between primary amino acid sequence and plastid targeting. Thus, it has proved difficult to predict plastid transit peptides (section 3.1.3.1).

Considerable insight into the workings of the plastid transit peptide has been gained from a recent study of deletion mutants of the *T. gondii* ferredoxin $NADP^+$ reductase leader.[273] Experiments were interpreted in the light of consensus motifs identified by a computational technique (TEIRESIAS Sequence Pattern Discovery[274]). Additionally, precise observations of intracellular protein localisation were achieved by deconvolution of micrographs. Hence it was possible to distinguish between targeting to

the plastid periphery (possibly synonymous with the periplastid space) and the plastid lumen.

By these methods, it was shown that the transit peptide possessed a linear microdomain architecture: the N-terminal of the transit peptide (probable motif V[V/S]SF) mediated escape from the ER after co-translational cleavage of the signal peptide; whereas, the central microdomain controlled trafficking to the plastid periphery. However, the master-targeting switch appeared to be the C-terminal of the plastid transit peptide. Absence of the C-terminal microdomain simultaneously dispatched the protein to the plastid periphery; parasitophorous vacuole; and rhoptries. Perhaps the C-terminal microdomain "grabs" the outer primary membrane of the plastid, thereby saving the protein from being "washed away."[236]

van Dooren, *et al.*[77] implicated PfSPP in the targeting process, and suggested proteolysis took place in the lumen of the plastid. Indeed, cleavage may be absolutely required for entry into the plastid lumen.[273] It has also been suggested that multiple cleavages occur. Furthermore, given the apparent lack of cleavage site consensus, it is tempting to speculate that several endopeptidases with different specificities might be involved. One final curiosity is that alternative splicing allows PfSPP to share a plastid-targeting leader with *P. falciparum* porphobilinogen synthase (PfPBGS; GenBank AY064477), an enzyme involved in haem biosynthesis.[77]

3.1.2.3. The third factor.

While it seems that a bipartite leader is necessary and sufficient for plastid targeting, precedent in plants suggests that accessory factors may enhance the process. For

example, Hsp70-like proteins can interact with chloroplast transit peptides,[275] thereby maintaining an unfolded state appropriate for passage through the chloroplast membranes. Foth, *et al.*[79] suggested the existence of a similar phenomenon in Apicomplexan plastid trafficking, citing the finding that DnaK-binding sites are abundant (> 90 % of leader sequences examined) in *P. falciparum* plastid transit peptides. Furthermore, ablation of the Hsp70-binding site of PfACP[L]-GFP directed a proportion of the fluorescent reporter to the parasitophorous vacuole;[79] while a similar experiment conducted in *T. gondii* abolished plastid-targeting altogether.[271] Further circumstantial evidence for DnaK involvement derives from the co-purification of an Hsp70-like chaperone (TgBiP) (and a subtilisin, TgSUB1) with TgL9[L]-GST fusions over-expressed in *T. gondii*.[276]

3.1.3. Identification of plastid-targeted proteins.

3.1.3.1. In silico *techniques.*

In spite of the obscure nature of the plastid transit peptide (section 3.1.2.2), *ab initio* predictions based on amino acid sequence have been made using an artificial neural-network (ANN) based algorithm (PATS).[78] Precedent (ChloroP) suggested that this might be a useful approach. However, as has been noted for ChloroP,[277] the problem with ANN-based algorithms is that it is difficult to know how they make decisions. This motivated the development of a rule-based algorithm (PlasmoAP) that was validated *in vivo* by selective mutation of PfACP[L].[79]

3.1.3.2. In vivo *techniques.*

In plants, it is possible to fractionate the chloroplast and perform a proteomic analysis. In spite of concerted efforts, such a technique has not yet been achieved in the Apicomplexan plastid. Therefore, localisation of plastid proteins relies entirely on immunocytochemistry (fluorescence microscopy or immunogold EM); or transient transfection. To my knowledge, only PfACP, PfPBGS[278] and (controversially[236]) PfFC[110] have been localised to the plastid by antibody treatment of fixed cells. Indeed, more common is to fuse the leader to a fluorescent reporter, e.g., GFP or *Discosoma* Red (DsRed).[267] Nevertheless, only a small number of proteins have been localised by these methods (Table 3.1).[97,99,108,110,271,278-280]

3.1.4. Chapter objective and overview.

The objective of this chapter is to describe preliminary steps in work intended to reveal the intracellular localisation of PfSufS. After the Materials and Methods, the results section begins with a discussion of the bioinformatic prediction of plastid localisation for PfSufS. Subsequently, transfection experiments in *P. falciparum* using the PfSufS leader fused to GFP are described. Transfection experiments in *T. gondii* using a synthetic *P. yoelii* SufS leader fused to a fluorescent reporter are also discussed. Finally, I report the development and initial evaluation of peptide antibodies intended for PfSufS immunolocalisation experiments.

TABLE 3.1.

PROTEINS DEMONSTRATED TO TARGET THE PLASTID *IN VIVO*.

Only published observations are reported. *Validity of localisation evidence has been challenged by Wilson.[236] Evidence codes are as follows. TF: transfection of cells with vector encoding a fluorescent reporter fused to the N-terminal leader sequence of the protein under investigation; IF: immunofluorescence with specific antibody directed against the protein under investigation; IG: immunogold electron microscopy with specific antibody directed against the protein under investigation.

Protein	Evidence	Organism	References
Acyl carrier protein	TF; IF; IG	*P. falciparum; T. gondii*	Waller, et al. [Ref. 97; 279]
Apicoplast dihydrolipoamide dehydrogenase	TF	*P. falciparum*	McMillan, et al. [Ref. 280]
β-ketoacyl-ACP synthase III	TF	*P. falciparum*	Waller, et al. [Ref. 279]
Ferrochelatase*	IG	*P. falciparum*	Varadharajan, et al. [Ref. 110]
Porphobilinogen synthase	IG	*P. falciparum*	Dhanasekaran, et al. [Ref. 278]
Pyruvate dehydrogenase E1α	TF	*P. falciparum*	Foth, et al. [Ref. 99]
Pyruvate dehydrogenase E2	TF	*P. falciparum*	Foth, et al. [Ref. 99]
Ribosomal protein S9	TF; IF	*T. gondii*	Yung, et al.; Waller, et al. [Ref. 271; 97]

3.2. MATERIALS AND METHODS.

3.2.1. Computational techniques.

Bipartite plastid-targeting sequence was predicted using the PATS[78]* and PlasmoAP[79]† algorithms. Signal sequence was predicted using SignalP 3.0;[289]‡ while mitochondrial targeting sequence was predicted using PlasMit.[290]¶ DnaK binding sites were predicted with the rudialg3.awk algorithm.[291]§ Quality of PCR primers was checked using the Amplify software. DNA sequences were assembled using the ABI Prism Sequencing Analysis software (Perkin Elmer) running on an Apple Power Mac G4. Molecular models were analysed with RasMol.

3.2.2. Parasite culture techniques.

P. falciparum 3D7 parasites were cultured by standard methods.[252] Parasites were cultured in leukocyte-depleted RBCs obtained from the National Blood Service (UK); or in whole blood obtained from healthy human volunteers (consistent with MRC ethical guidelines). Briefly, *ca.* 10 mL whole blood was recovered by venipuncture into a tube containing 2.8 mL sterile anticoagulant citrate dextrose (2.45 g glucose, 2.20 g tri-sodium citrate and 0.89 g citric acid in 100 mL redistilled water). After gentle mixing, the tube was placed at 4 °C to allow the cells to separate from the serum by gravity. Serum was aspirated and packed cells used in transfection experiments for < 28 d. *T. gondii* tachyzoites were grown in MDCK cells by standard techniques.[286]

---

* http://gecco.org.chemie.uni-frankfurt.de/pats/pats-index.php
† http://plasmodb.org/restricted/PlasmoAPcgi.shtml
‡ http://www.cbs.dtu.dk/services/SignalP/
¶ http://gecco.org.chemie.uni-frankfurt.de/plasmit/index.html
§ http://www.bijvoet-center.nl/cpc/labmembers/Stefan/DnaKbinding

3.2.3. Nucleic acid techniques.

Nucleic acid techniques were as described (chapter 2). Primer sequences are indicated in Table 3.2.

3.2.4. Transfection experiments in *Plasmodium falciparum*.

3.2.4.1. P. falciparum *transfection constructs.*

Two vectors were used for transient transfection of *P. falciparum*: pHH2 and pSSPF2. The pHH2 vector was created by Waller, *et al.*[279] and encodes GFP-PfACP[L]. The vector also encodes mutant human dihydrofolate reductase (hDHFR), allowing selection with the drug WR99210. The pSSPF2 vector is a modified version of pHH2 developed by Sato, *et al.*[223] Vector pHH2 was a gift from Dr. A. F. Cowman (WEHI, Melbourne, Australia); pSSPF2 was provided by Dr. S. Sato (NIMR).

3.2.4.2. Cloning of P. falciparum *SufS leader sequences.*

Exon 1 (PfsufS[E1]) and sequence encoding the first 54 (PfsufS[L_{54}]) and 72 (PfsufS[L_{72}]) amino acids of PfSufS were separately amplified by RT-PCR from *P. falciparum* 3D7 cDNA. PfsufS[E1] was amplified with primers 3.1 and 3.2. Thermocycling conditions were: denaturation at 94 °C for 1 min; followed by 35 cycles of 98 °C for 10 s; 45 °C for 2 min; and 72 °C for 10 min.

TABLE 3.2.

PCR PRIMERS USED IN CHAPTER 3.

ID	DNA sequence
3.1	5'AGATCTAATGTTAAGAGGCCCTAGATGTCTC
3.2	3'CCTAGGTTATAAAACTCACTCATTTTCTCTAT
3.3	5'GGGAGATCTAAAATGTTAAGAGGCCCTAGATGTCTCTACAT
3.4	3'GGGCCTAGGCTTTTCATCTTTTGTCAATC
3.5	3'GGGCCTAGGAAAATGTTCTCTTACATTTTTAAAATAATC
3.6	5'AGATCTAAGATGAACAACAAAAGTATTTGCATCTTGCTGC
3.7	5'TGCTCTTCCTTAAGATTACAGGATACTACAGTTACTACAT
3.8	5'TAGCCCATTGAAAAAGAACATTTATACACACCCAATCTAC
3.9	5'AAGAGACATATCCTCAAGTTGCAAAATGAAAAACCTGAGA
3.10	5'TCGACTCGAACATTATTAACTACTTTAAGAACATTCGCTC
3.11	5'TGATTTTCCCTTCTTTCAACAGAATAATTCACCCATCTAC
3.12	5'TTTGATAATGCTGCAACTACACATAAACCTGCTTCTGTAA
3.13	3'CCTAGGTTTGATCTTTTCAATTACAGAAGCAGGTTTATGT
3.14	3'GTAGTTGCAGCATTATCAAAGTAGATGGGTGAATTATTCT
3.15	3'GTTGAAAGAAGGGAAAATCAGAGCGAATGTTCTTAAAGTA
3.16	3'GTTAATAATGTTCGAGTCGATCTCAGGTTTTTCATTTTGC
3.17	3'AACTTGAGGATATGTCTCTTGTAGATTGGGTGTGTATAAA
3.18	3'TGTTCTTTTTCAATGGGCTAATGTAGTAACTGTAGTATCC
3.19	3'TGTAATCTTAAGGAAGAGCAGCAGCAAGATGCAAATACTT

In order to amplify PfsufS[L_{54}], primers 3.3 and 3.4 were used; equivalents for PfsufS[L_{72}] were 3.3 and 3.5. Thermocycling conditions were: denaturation at 94 °C for 1 min; followed by 35 cycles of 98 °C for 20 s; 56 °C for 30 s; and 72 °C for 1 min.

PfACP[L] was excised from the parent vector by restriction digest (*Bgl*II and *Bln*I) and replaced with PfsufS[E1], giving the plasmid pHH2:PfsufS[E1]-GFP. PfsufS[E1] was additionally cloned into pSSPF2 between *Bgl*II and *Bln*I sites, giving rise to pSSPF2:PfsufS[E1]-GFP. PCR products of PfsufS[L_{54}] and PfsufS[L_{72}] were digested with *Bgl*II and *Bln*I and inserted between these sites in pSSPF2, giving constructs pSSPF2:PfsufS[L_{54}]-GFP and pSSPF2:PfsufS[L_{72}]-GFP, respectively. Inserts were confirmed by restriction digest and Sanger sequencing. pHH2:PfACP[L]-GFP and pSSPF2:PfiscS[L]-GFP were used as positive controls. GFP is targeted to the plastid by PfACP[L][279] and to the mitochondrion by PfIscS[L].[223]

3.2.4.3. Preparation of DNA for transfection.

To prepare DNA for transfection, INVαF′ cells harbouring vector were grown for 12 h at 37 °C in 100-500 mL Luria-Bertani broth (10 g bactotryptone, 5 g yeast extract, 10 g NaCl per 1 L redistilled water) with 0.2 % D-glucose and 100 $\mu g.mL^{-1}$ ampicillin. Plasmid DNA was purified by Wizard Plus Maxiprep (Promega) or QIAfilter Plasmid Maxi (Qiagen). DNA was ethanol precipitated and the concentration and purity assayed by 260/280 ratio and restriction digest. Immediately prior to transfection, DNA (100 μg) was dissolved in 50 μL of sterile TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA).

3.2.4.4. Preparation of parasites for transfection.

Two days before electroporation, *P. falciparum* erythrocytic stages were synchronised in 10 % D-sorbitol by standard methods.[292] This ensured that the culture was predominantly at the ring stage for electroporation.

3.2.4.5. Electroporation of parasites.

Cells were loaded with DNA by electroporation as follows. Parasites (1 % parasitaemia, 200 μL packed cell volume) were re-suspended in 200 μL sterile cytomix (120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2HPO_4/KH_2PO_4, 25 mM HEPES, pH 7.6). DNA in TE was added and the suspension mixed by gentle pipetting. The suspension was placed in a 0.2 cm gap electroporation cuvette (Biorad Genepulser) and electroporated on a BTX Electro Cell Manipulator 600 (parameters were set to 310 V, 950 μF, 24 Ω). For all electroporations, the recorded voltage was 290-300 V and the pulse time was 8-10 ms.

3.2.4.6. Recovery and drug selection.

Electroporated cells were transferred to 10 or 20 mL RPMI-1640 with Albumax (Invitrogen), 1 % glutamine and supplemented with 2 % AB+ human serum (National Blood Service, UK) and incubated under 5 % O_2, 5 % CO_2 in N_2 at 37 °C. After 2 d recovery, 2.5-10 nM of the anti-folate drug WR99210 (Jacobus Pharmaceutical, Princeton, NJ, USA) was added in RPMI-1640. Gas, medium and drug were changed regularly. The health of the culture was checked by thin blood film stained with

Giemsa's stain. To detect fluorescent parasites, live cultures were viewed by UV microscopy (Zeiss).

3.2.4.7. Growth assay with [8-^{3}H]hypoxanthine.

Electroporated cells were transferred to 96-well microtitre plates. A volume (1 μL) of [8-^{3}H]hypoxanthine (Amersham) was added to each well. After incubation for 48 h, parasites were harvested onto Wallac filtermats using a Tomtec cell harvester. Filtermats were soaked in 20 mL Betaplate Scint liquid scintillation cocktail (Perkin Elmer) and counted on a Wallac 1205 Betaplate liquid scintillation counter.

3.2.5. *Toxoplasma gondii* experiments.

3.2.5.1. Construction of a synthetic SufS leader.

The approach of Jomaa, *et al.*[107] was used. The first 100 amino acids of the *P. yoelii* SufS N-terminal were predicted to encode the leader sequence. The leader was back-translated with codon bias harmonised for *T. gondii* (by reference to the Codon Usage Database*) and flanked with restriction sites (*Bgl*II and *Bln*I). Harmonised sequence was divided into fifteen 20-mers and codon usage further adjusted to ensure melting temperatures of 52-56 °C. The 20-mers were concatenated into seven overlapping 40-mers and complimentary sequence was generated, giving a total of 14 oligomers (primers 3.6-3.19) to cover synthesis of the leader.

* http://www.kazusa.or.jp/codon/

Synthetic PySufS leader (PySufS[L_s]) was assembled by recursive PCR essentially as described.[293] Briefly, equal volumes of the 14 oligomers were combined and diluted 1/5 into a PCR reaction mixture containing 1 × *Pfu* buffer (Promega), 5 mM $MgSO_4$, 1 mM of each dNTP and 3 U *Pfu* DNA polymerase (Promega). The final concentration of each oligomer in the reaction mixture was 180 nM. Oligomers were assembled by denaturation at 94 °C for 1 min followed by 24 cycles of 94 °C for 30 s, 50 °C for 30 s and 72 °C for 2 min. A second PCR was conducted to amplify full length product using the extreme 5′ and 3′ primers (3.6 and 3.13). Thermocycling conditions were: 94 °C for 1 min, followed by 24 cycles of 94 °C for 45 s, 68 °C for 45 s and 72 °C for 2 min. The amplicon was inserted into the pCR-Script Amp SK(+) plasmid (Stratagene) by blunt end cloning. Insert was confirmed by restriction digest and Sanger sequencing.

3.2.5.2. T. gondii *targeting constructs and transfection.*

Two similar constructs were used for transient transfection of *T. gondii*: (1) pSAG1/CAT/TUB1/TgACP leader-GFP; and (2) pSAG1/CAT/TUB1/TgACP leader-DsRed. These vectors were originally developed by Soldati and Boothroyd[288] and encode *T. gondii* major surface antigen (SAG1), chloramphenicol acetyltransferase (CAT) and α-tubulin (TUB1). Additionally, the vectors encode a fluorescent marker [green fluorescent protein (GFP) or *Discosoma* red (DsRed)] in frame to the *T. gondii* acyl carrier protein leader sequence (TgACP[L]). TgACP[L] targets the fluorescent marker to the plastid organelle. Both vectors were a gift from Dr. S. Sato.

TgACP[L] was excised from the vectors by restriction digest (*BglII* and *BlnI*) and replaced with PySufS[L_s]. Insert was confirmed by restriction digest and Sanger sequencing. DNA was prepared for transfection as described (section 3.2.4.3).

Immediately prior to transfection, DNA (25 pmol per transfection) was dissolved in 100 μL of sterile cytomix. Transfections were performed by Dr. R. J. M. Wilson (NIMR), as described.[288] To detect fluorescent parasites, live cultures were viewed under a fluorescence microscope (Zeiss).

3.2.6. Design, production and evaluation of peptide antibodies.

3.2.6.1. Production of peptides and polyclonal sera.

Antibodies were prepared by Sigma-Genosys (UK). Peptides from PfSufS were synthesised via solid-phase Fmoc chemistry validated by mass spectrometry (MS) and reversed phase high performance liquid chromatography (HPLC); and conjugated to Keyhole limpet haemocyanin ([KLH]). Peptides were: HIKYDIKKMKKNKSC-[KLH] (BURN-1); and [KLH]-[C]KYYLKVPDTSRIS (BURN-2). Immunizations were done in New Zealand White rabbits according to IACUC animal welfare guidelines. The immunization regimen lasted 77 d and comprised one priming immunization (200 μg) in complete Freund's adjuvant; followed by five boosting immunizations (100 μg) in incomplete Freund's adjuvant. Serum was used for Western blotting without purification.

3.2.6.2. Screening of polyclonal anti-sera.

Polyclonal anti-sera were screened by Western blotting against *P. falciparum* 3D7 lysate or recombinant PfSufS proteins (chapter 4). *P. falciparum* 3D7 lysate was prepared from mixed erythrocytic stages (parasitaemia *ca.* 10 %; packed cell volume *ca.*

2 mL) by repeated freezing in a dry ice/ethanol slush; and thawing in a 37 °C water bath. Lysed parasites were diluted in PBS; mixed with reducing SDS loading buffer (NEB); and heated to 70 °C for 10 min prior to separation through 12 % BT NuPAGE gel (Invitrogen). Gels were transferred to nitrocellulose membrane at 20 mA for 15 h; blots were blocked for 3 h in 5 % w/v BSA fraction V (Sigma), 0.1 % v/v Tween 20 (Sigma) in PBS. Blots were immersed for 1 h in serum diluted 1:1000 in blocking buffer followed by incubation with 1:3000 goat anti-rabbit IgG (H+L)-HRP conjugate (Bio-Rad). Cross-reactive bands were revealed by ECL (Amersham) and imaged using Amersham Hyperfilm. Images were captured using an Epson Perfection 2480 Photo scanner and molecular weights of proteins estimated by comparison to Mark12 unstained standards (Invitrogen) using image analysis software (UVP LabWorks).

3.3. RESULTS AND DISCUSSION.

3.3.1. Structure of the bipartite leader of *Plasmodium falciparum* SufS.

3.3.1.1. Plastid-targeting properties of PfSufS[L].

To predict the intracellular localisation of PfSufS, the strict decision tree employed by Ralph, *et al.*[80] was applied. Submission of the PfSufS protein sequence to PATS (p = 0.947) and PlasmoAP (5 of 5 tests positive) indicated the presence of a bipartite plastid-targeting signal. It is also worth noting that PfSufS is non-mitochondrial, as shown by PlasMit (p = 0.99).

3.3.1.2. Signal peptide.

Further inspection of the N-terminal portion of PfSufS was done using the SignalP neural network (NN) and Hidden Markov Model (HMM) algorithms. SignalP-NN predicted a signal peptide (5/5 tests positive) while SignalP-HMM favoured an un-cleaved signal anchor (p = 0.501) over a signal peptide (p = 0.495). Both algorithms predicted signal peptidase cleavage between C22 and Y23, thus delimiting the signal peptide to the first 22 amino acids of PfSufS. Based on this prediction, the plastid transit peptide would begin at Y23.

3.3.1.3. Plastid transit peptide.

Further analysis was conducted upon the predicted plastid transit peptide, starting at Y23 and predicted to end at I60. Consistent with this, alignment with bacterial SufS sequences (lacking intracellular targeting signals) suggested that the mature PfSufS enzyme started at about D62. Pair-wise comparison with the plastid transit peptide of PfACP indicated significant divergence at the amino acid level, although both regions were predicted to carry a net positive charge and were enriched for arginine (data not shown).

It has been suggested that binding of DnaK may be an additional determinant of plastid-targeting in the Apicomplexa, mimicking the situation of chloroplast transit peptides.[275] An algorithm developed for prokaryotic sequences is useful for detection of DnaK binding sites in *P. falciparum* transit peptides.[79] Indeed, the algorithm revealed a possible DnaK binding site ($\Delta\Delta G_K$ = -11.44 kJ.mol^{-1}) centred on residue Y33 of PfSufS.

3.3.1.4. Model for plastid trafficking of PfSufS.

The predictions summarised above reasonably suggested that PfSufS targets the plastid. Experimental methods have been applied to verify such predictions (sections 3.3.2-3.3.4).

My prediction is that full-length PfSufS (65 kDa) is expressed on cytosolic ribosomes and then cleaved by a signal peptidase, giving an incompletely processed product of ~ 62 kDa. Further clipping by a stromal processing peptidase (e.g., PfSPP) would occur in the plastid, leaving a mature enzyme of ~ 57 kDa. The predicted structure of PfSufS[L] is summarised in Fig. 3.1.

FIGURE 3.1.

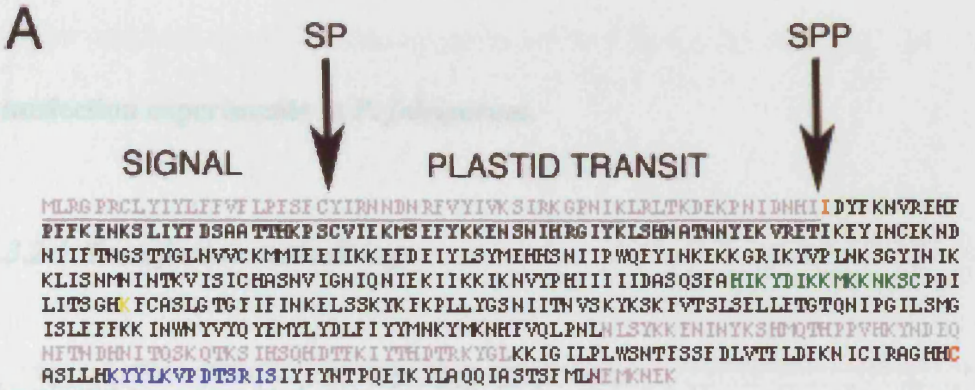
IMPORTANT TARGETING FEATURES OF *PLASMODIUM FALCIPARUM* SUFS.

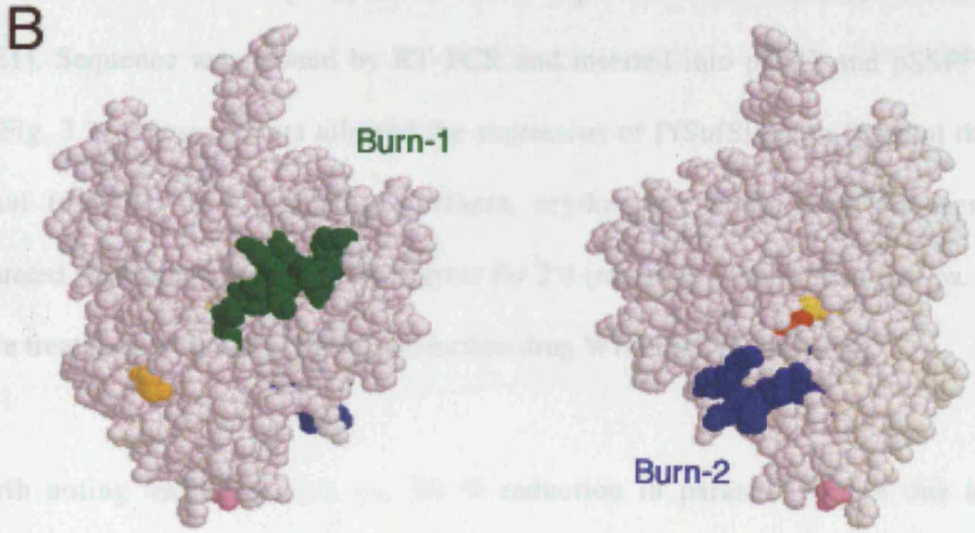
A. Major features of PfSufS protein sequence.

The signal peptide is predicted to run M1-C23, with signal peptidase (SP) cleavage occurring between C23 and Y24. The plastid transit peptide may stretch from Y24-I60, with cleavage by stromal processing peptidase (SPP) possibly occurring between I60 and D62. As determined by alignment with EcSufS, mature PfSufS runs D62-K546. Peptides used for generation of peptide antibody are marked in green (Burn-1) and blue (Burn-2). The putative PLP-binding lysine (K291) is marked in yellow while the active site cysteine (C497) is marked in red. Regions excluded from the homology model are marked in grey, while the extremes of the modelled region are marked in orange (N-terminal) and purple (C-terminal). Bioinformatic methods are described in section 3.2.1.

B. Homology model of PfSufS showing position of peptides.

Peptides used for generation of peptide antibody are marked in green (Burn-1) and blue (Burn-2). The N-terminal is marked orange while the C-terminal is marked purple. The putative PLP-binding lysine (K291) is marked in yellow while the active site cysteine (C497) is marked in red. The molecule in the left panel is a 90° rotation in the x-plane of the molecule in the right panel.


A
SP
SPP
SIGNAL
PLASTID TRANSIT


B
Burn-1
Burn-2

It is tempting to speculate that cofactor (PLP) binds PfSufS apoenzyme in the plastid lumen. However, it is worth noting that a protein expected to be critical for this process, pyridoxal kinase (PfPdxK; GenBank NP_703820),[281] resides in the cytosol in *P. falciparum*.[282] Dimerization of PfSufS; and complex formation with other components of the SUF system (e.g., PfSufBCDE) may additionally take place in the lumen of the plastid.

3.3.2. Transfection experiments in *P. falciparum*.

3.3.2.1. Transfection methodology.

To test the plastid-targeting prediction for PfSufS (section 3.3.1), transient transfection experiments were done in *P. falciparum*. Initial experiments were conducted with PfSufS[E1]. Sequence was cloned by RT-PCR and inserted into pHH2 and pSSPF2 vectors (Fig. 3.2). These vectors allowed the expression of PfSufS[E1] to fused at the C-terminal to GFP. In a typical experiment, erythrocytic stage parasites were electroporated with vector and allowed to grow for 2 d (reaching a parasitaemia of *ca.* 4 %), before treatment with the anti-folate selection drug WR99210 (2.5-20 nM).

It is worth noting that there was *ca.* 50 % reduction in parasite growth due to electroporation, as assessed by measurement of [^{3}H]hypoxanthine uptake over the 48 h post shock period (Fig. 3.3). In addition, treatment with drug was rapidly parasiticidal and *ca.* 4 d after first exposure no live parasites were visible after extensive (*ca.* 20 min) searching of Giemsa-stained blood films.

FIGURE 3.2.

RESTRICTION DIGEST OF PFSUFS[E1]-GFP CONTRUCTS.

Vector DNA of pHH2:PfsufS[E1]-GFP (A) and pSSPF2:PfsufS[E1]-GFP (B) was isolated by maxiprep and digested with *BglII* and *BlnI*, thus excising PfsufS[E1]. Digests were resolved through 1 % agarose gel and viewed under ultraviolet light after treatment with ethidium bromide. Abbreviations are as follows. bp: base pair; M: 123 bp ladder (Sigma); D: digest.

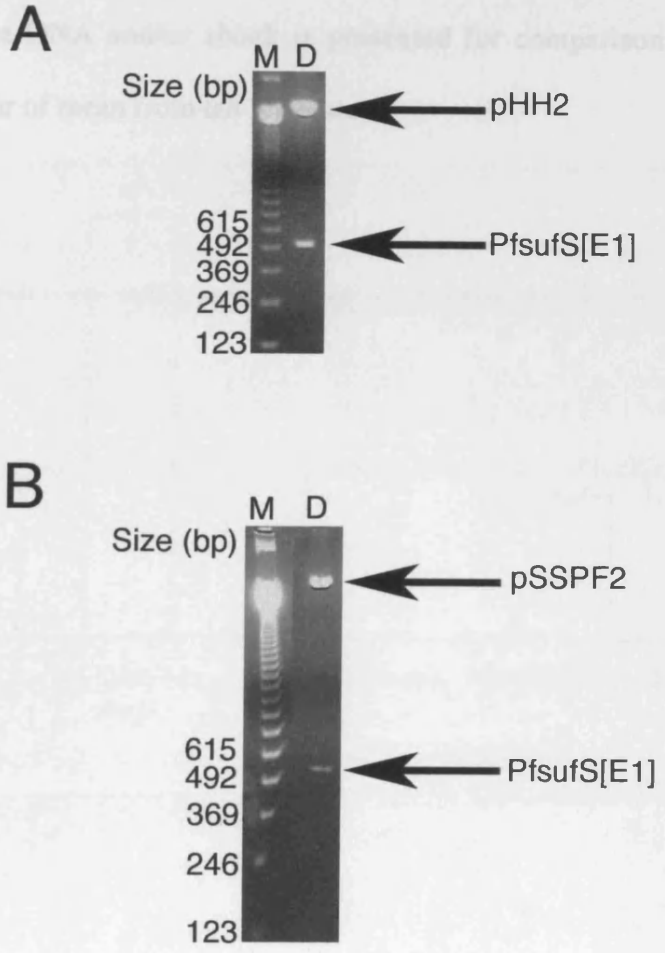
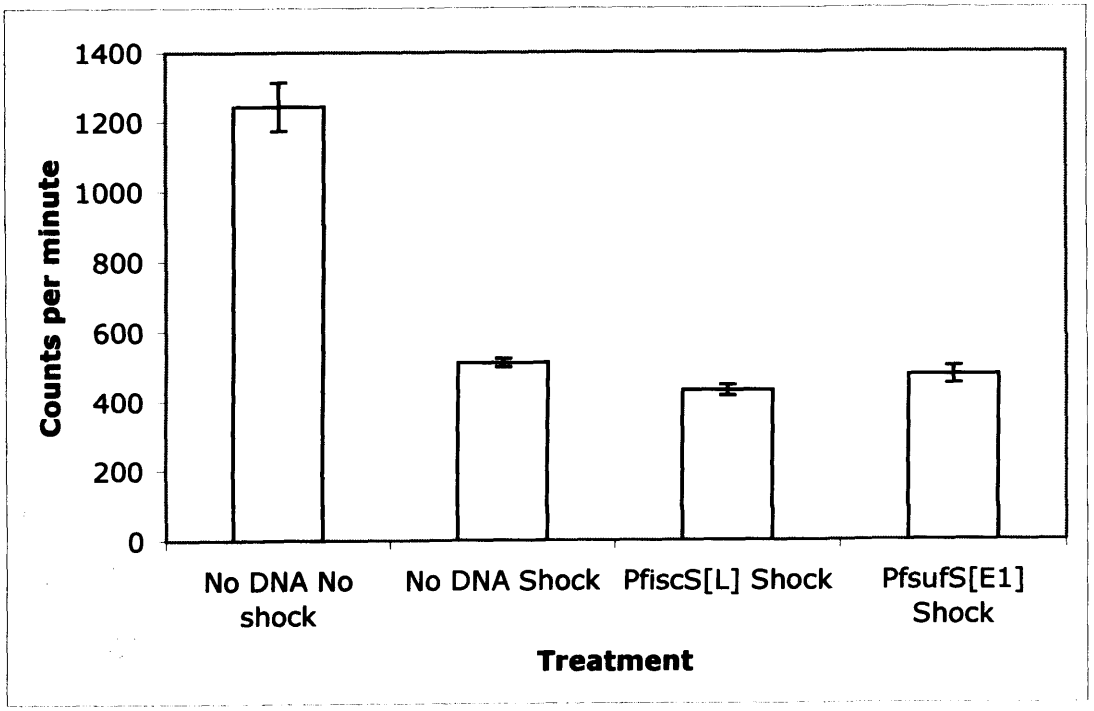

A
M D
Size (bp)
pHH2
615
492
369
246
123
PfsufS[E1]
B
M D
Size (bp)
pSSPF2
615
492
369
246
123
PfsufS[E1]

FIGURE 3.3.

EFFECT OF ELECTROPORATION ON THE GROWTH OF *PLASMODIUM FALCIPARUM*.

Electroporation of *P. falciparum* was done with pSSPF2:PfiscS[L]-GFP or pSSPF2:PfsufS[E1]-GFP constructs as described (section 3.2.4.5). Growth of cells for 48 h after electroporation was monitored by uptake of [^{3}H]hypoxanthine. Growth of cells without DNA and/or shock is presented for comparison. Error bars represent standard error of mean from ten replicates.


1400
1200
1000
800
600
400
200
0
Counts per minute
No DNA No shock
No DNA Shock
PfiscS[L] Shock
PfsufS[E1] Shock
Treatment

Following drug selection, it was assumed that a small population of parasites would persist that harboured the transfection vector and thus were resistant to drug.

To encourage the expansion of this population, a healthy haematocrit (1-5 %) was maintained by adding fresh RBC and changing medium and drug regularly. This is known as the recovery period and lasted several weeks. Throughout, the culture was screened for the appearance of parasites.

3.3.2.2. Experiments with vector pHH2.

Initial experiments were done with pHH2:PfsufS[E1]-GFP and pHH2:PfACP[L]-GFP was used as a positive control. However, in several independent experiments, transfectants did not arise. This indicated problems with the transfection and culturing technique. To investigate this, I adopted a systematic approach to account for the lack of transfectants.

First, the quality of the DNA used for the transfection was considered, as assessed by 260/280 ratio and restriction digest. Preparations of vector DNA were performed from small culture volumes (100-500 mL). This is because overloading of maxiprep kits can contaminate the DNA with bacterial products (e.g., polysaccharides). Phenol-chloroform extraction was avoided during DNA preparation, because traces of phenol are toxic to parasites. Repeated washing in 70 % ethanol during ethanol precipitation steps reduced salt contamination. After taking these measures, DNA preparations were extremely pure. Ratios 260/280 were typically 1.8 and DNA was cleanly digested by restriction enzymes.

Second, the conditions in the electroporation cuvette were examined. To do this, transfections were done at various parasitaemias (0.5-2 %) and haematocrits (25-50 %). It was decided to leave electrical conditions unchanged for two reasons: (1) they have been extensively evaluated (Sato, unpublished observations) and (2) electroporation produced little cell lysis in my experiments.

Third, the environment inhabited by the parasites during the recovery period was examined. After transfection, it was assumed that parasites were exquisitely sensitive to environmental conditions. One key condition was the concentration of WR99210. Concentrations between 2.5-20 nM were tested in independent transfections. Also examined was the amount of physical space occupied by the cells. Cells in a confined area are more accessible to secreted growth factors that may facilitate recovery. Hence, flasks were incubated either upright (close cell packing) or on the side (loose cell packing). Another factor of importance may be mechanical disturbance of the cells. In one regimen, the flask was gassed every day and the medium and drug changed every other day. Therefore, to minimise disturbance, the regimen was altered to a gas, medium and drug change every 5 d. Other measures were taken to reduce mechanical stress, for example, by avoiding centrifuging the cultures.

In spite of these efforts, no transfected parasites were observed. Parasites transfected with pHH2 can take $\leq$ 55 d after electroporation to appear.[283] The longest time I have run an experiment is 30 d. This was because of significant red blood cell lysis (i.e. loss of > 25 % haematocrit) at > 30 d. Even with regular addition of fresh RBC, the bulk of the culture dates from the time of electroporation. Thirty days may approach the maximum survival time of RBC (value *in vivo* is 110 ± 40 d[284]).

3.3.2.3. Experiments with vector pSSPF2.

Given the lack of immediate success with pHH2, I tried pSSPF2, an improved vector recently developed by Dr. S. Sato. To this end, PfsufS[E1] was excised from pHH2 and cloned into pSSPF2. The performance of two different maxiprep kits was compared for the preparation of this construct. In a typical experiment, the Wizard Maxiprep (Promega) yielded *ca.* 70 μg of vector per 100 mL of culture, while the QIAfilter Plasmid Maxi (Qiagen) kit yielded *ca.* 110 μg of vector per 100 mL of culture. One disadvantage often attributed to pHH2 (as compared to pSSPF2) is a lower yield of vector DNA from maxiprep. As both vectors were available containing an identical construct (PfsufS[E1]-GFP), this observation was tested. In a typical maxiprep, the yield of pHH2:PfsufS[E1]-GFP was < 30 μg vector per 100 mL of culture, while for pSSPF2:PfsufS[E1]-GFP, it was 40-110 μg vector per 100 mL of culture. pHH2 preparations were often contaminated with bacterial polysaccharides.

Construct pSSPF2:PfiscS[L]-GFP was successfully used in this laboratory to target GFP to the mitochondrion in *P. falciparum*[223] and was thus used as a positive control. However, in transfection experiments with pSSPF2 constructs, no parasites appeared in either positive control or treatment. A parallel experiment conducted by Dr. S. Sato using pSSPF2:PfsufS[E1]-GFP lead to no observable fluorescence in parasites. Although, experience of other workers (Drs. S. Sato, S. J. Lucas and M. Kadekoppala, NIMR, pers. comm.) indicated that repeated experiments are often necessary, I decided to re-engineer the construct.

3.3.2.4. Re-engineering of the pSSPF2 construct.

It was noted that PfsufS[E1] encoded a proportion of mature enzyme sequence; hence it probably exceeded the length of the leader. Therefore, shorter constructs encoding the first 54 (pSSPF2:PfsufS[L_{54}]-GFP) and 72 (pSSPF2:PfsufS[L_{72}]-GFP) amino acids of PfSufS genetically fused to GFP were generated. The first 54 and 72 amino acids were chosen in order to bracket the predicted leader sequence (about 60 amino acids) and for convenience of PCR primer design. I noted that a consensus sequence (AAAA) is commonly found immediately 5′ to the start codon of *P. falciparum* ORFs.[285] Hence, AAA was incorporated into the forward primer for both constructs. Other workers will continue experiments with these constructs.

3.3.3. Transfection experiments in *Toxoplasma gondii.*

Transient transfection experiments were also attempted in *T. gondii.* As compared to *P. falciparum*, *T. gondii* is readily transfected[286] and has clear sub cellular morphology.[287] However, the *T. gondii* SufS leader (TgSufS[L]) could not easily be deduced from the incomplete ToxoDB database. Identification of TgSufS[L] by inspection of genomic reads was especially difficult, because *T. gondii* sequence often has numerous short introns that are difficult to recognise. Several *ab initio* gene models of TgSufS (e.g., TgTwinScan_1344, TgGlmHMM_1221 and TgTigrScan_3457) are so confused by this phenomenon as to be of little value. Predictably, ESTs are only available for the C-terminal of the protein (chapter 2).

Therefore, the strategy of Jomaa, *et al.*[107] was adopted. Specifically, the predicted leader sequence of *P. yoelii* SufS (PySufS[L_s]) was synthesised *de novo* by recursive PCR,

with codon usage harmonised for *T. gondii* (Fig. 3.4). PySufS was chosen as the leader sequence is less divergent from TgSufS than PfSufS (Dr B. Clough, pers. comm.) The synthetic leader was cloned into the transfection vectors pSAG1/CAT/TUB1/GFP and pSAG1/CAT/TUB1/DsRed. Thirty separate clones were examined before correct sequence was identified, indicating the low success rate of recursive PCR.

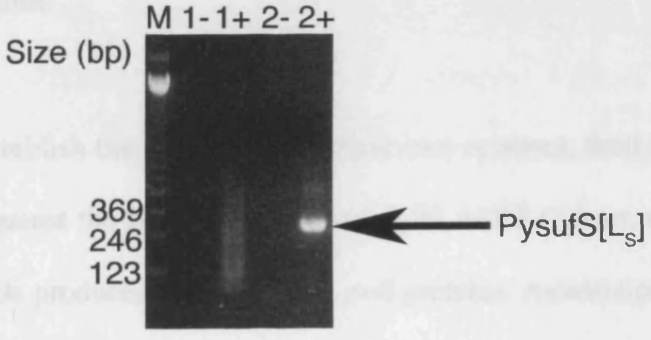
Parasites were transfected and the vector pSAG1/CAT/TUB1:TgACP[L]-GFP was used successfully as a positive control. However, no fluorescence was detectable following transfection with the PysufS[L_s]-GFP or PysufS[L_s]-DsRed constructs.

It is possible that the PysufS[L_s] fusion proteins are ejected from the cell (where fluorescence would be undetectable). This is because plastid transit peptides are poorly conserved and the *P. yoelii* version is probably not recognised by *T. gondii*. Hence the first part of the synthetic leader (signal peptide) is targeting the fluorescent tag to the secretory pathway. It might be useful to measure CAT activity in cells transfected with the pSAG1/CAT/TUB1 constructs, to definitively confirm vector uptake. CAT is encoded by the vector and enzymatic activity is considered to be a measure of successful transfection.[288]

FIGURE 3.4.

PCR SYNTHESIS OF PYSUFS[L_S].

PysufS[L_s] was assembled by recursive PCR (section 3.2.5.1). Briefly, 14 overlapping primers encoding PySufS[L_s] were assembled by PCR (1+). Full-length product was resolved in a second round of PCR directed by the extreme 5′ and 3′ primers (2+). Negative controls lacking DNA are presented (1−; 2−). Additional abbreviations are as follows. bp: base pair; M: 123 bp ladder (Sigma).


M 1- 1+ 2- 2+
Size (bp)
369
246
123
PysufS[L$_S$]

3.3.4. Development of peptide antibodies against *Plasmodium falciparum* SufS.

3.3.4.1. Selection of peptides.

PfSufS protein sequence was submitted to Sigma-Genosys in order to select epitopes that were: (1) hydrophilic; (2) antigenic; (3) suitable for Fmoc synthesis; and (4) readily conjugated to KLH. Only five peptides were suitable: two mapped to the leader sequence and were discarded immediately on the assumption that the leader was cleaved during entry to the plastid. Of the remaining three peptides, only two were deemed of suitable length (> 10 residues). Both mapped to the surface of the PfSufS homology model (Fig. 3.1) and thus were assumed to be accessible to antibody under native conditions.

In order to establish the uniqueness of these two epitopes, their protein sequences were scrutinised against the NCBI database using BLASTP ("short, nearly exact matches"). Neither peptide produced hits against *E. coli* proteins. As anticipated, the only strong hit (E = 2e-04) was against PfSufS; although weaker hits were observed against SufS from rodent malarias (*P. yoelii*, *P. chabaudi* and *P. berghei*). It is worth noting the lack of similarity to PfIscS. I decided that every bioinformatic precaution had been taken to ensure antibody cross-reactive to these peptides would be specific for PfSufS. Hence, the two peptides were synthesised; coupled to KLH; and used to immunize to New Zealand white rabbits.

3.3.4.2. Screening of peptide antibodies.

Serum samples were obtained from the rabbits at intervals of 5 d over the 77 d immunization protocol. In order to test if the sera exhibited specific cross-reactivity against PfSufS, Western blots were performed. The results of preliminary experiments will now be described. Initially, GST-PfSufS[M]χ (chapter 5) was used as a target. Both anti-Burn-1 and anti-Burn-2 serum decorated GST-PfSufS[M]χ (Fig. 3.5).

Dilutions of *P. falciparum* lysate were probed with anti-Burn-1 serum (Fig. 3.6). Both anti-Burn-1 and pre-immune sera cross-reacted with bands at 55 kDa, 44 kDa, 40 kDa and 31 kDa. However, anti-Burn-1 decorated one additional high molecular weight band (68 kDa) that could correspond to PfSufS (predicted size 65 kDa). There is considerable suspicion that the observed patterns on Western blots were artefactual. Thus, these data are of little value until confirmed. Furthermore, anti-Burn-2 needs to be investigated.

3.3.4.3. Future work.

In order to further investigate the cross-reactivity observed against *P. falciparum* lysate, it will be necessary to unequivocally demonstrate specificity of anti-Burn-1 serum for PfSufS. This will be most rapidly achieved by Western blotting against pure recombinant PfSufS[M]χ, lacking an affinity tag. Efforts to make such protein are discussed in chapter 5. Furthermore, it may be useful to generate monoclonal antibodies against the entire protein. This could be achieved by immunization of mice with pure NusA-PfSufS[M]χ (chapter 5).

FIGURE 3.5.

CROSS-REACTIVITY OF ANTI-PFSUFS SERUM AGAINST GST-PFSUFS[M]χ.

Western blotting was used to screen antisera generated by immunisation of rabbits with PfSufS peptides (Burn-1 and Burn-2). The target was a recombinant GST fusion of PfSufS (GST-PfSufS[M]χ) generated in chapter 5. Cross-reactivity of the recombinant protein with pre-immune serum is presented for comparison (Pre-im1; Pre-im2). Additional abbreviations are as follows. M: Mark12 unstained standards; S: partially purified GST-PfSufS[M]χ; Coom: Coomassie stained gel.

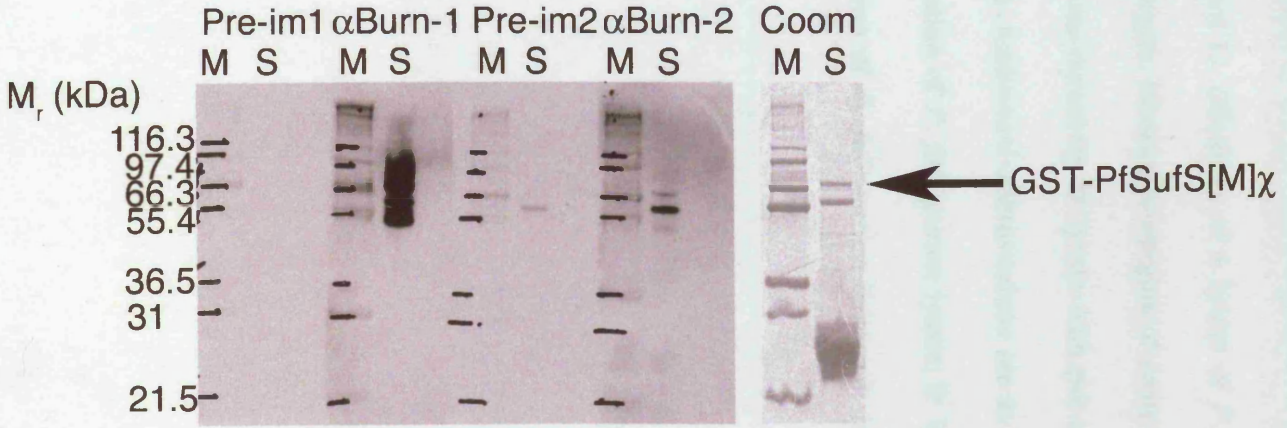
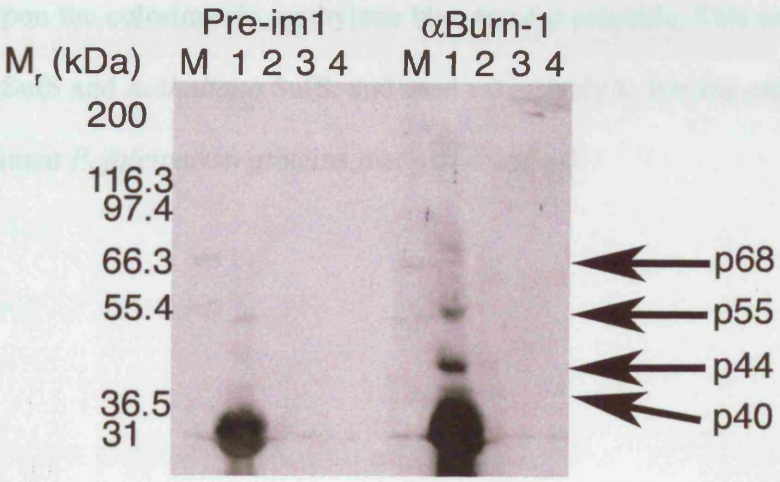

Pre-im1
αBurn-1
Pre-im2
αBurn-2
Coom
M S M S M S M S M S
M$_r$ (kDa)
116.3
97.4
66.3
55.4
36.5
31
21.5
GST-PfSufS[M]χ

FIGURE 3.6.

CROSS-REACTIVITY OF ANTI-PFSUFS SERUM AGAINST *PLASMODIUM FALCIPARUM* LYSATE.

Western blotting was used to screen antisera generated by immunisation of rabbits with a PfSufS peptide (Burn-1). Dilutions of a lysate of *P. falciparum* 3D7 erythrocytic stages were used as targets. Molecular weights of cross-reactive bands discussed in the text are indicated. Cross-reactivity of lysate with pre-immune serum is presented for comparison (Pre-im1). Additional abbreviations are as follows. M: Mark12 unstained standards; 1: 1/10 dilution of *P. falciparum* lysate; 2: 1/100: dilution of *P. falciparum* lysate; 3: 1/250 dilution of *P. falciparum* lysate; 4: 1/500 dilution of *P. falciparum* lysate.


Pre-im1
αBurn-1
M r (kDa)
M 1 2 3 4
M 1 2 3 4
200
116.3
97.4
66.3
55.4
36.5
31
p68
p55
p44
p40

CHAPTER 4.

DEVELOPMENT OF A COLORIMETRIC CYSTEINE DESULPHURASE ASSAY.

Plasmodium falciparum SufS is predicted to be a pyridoxal 5'-phosphate (PLP)-dependent cysteine desulphurase. Therefore, this chapter begins with a brief overview of PLP chemistry and PLP-dependent enzymes, followed by a detailed description of the reaction mechanism of *E. coli* SufS (EcSufS). Knowledge of bacterial cysteine desulphurases allowed the selection of a robust assay for cysteine desulphurase activity based upon the colorimetric methylene blue test for sulphide. This assay was validated with EcSufS and *A. thaliana* SufS; and used extensively to test the enzymatic activity of recombinant *P. falciparum* proteins made in chapter 5.

4.1. INTRODUCTION.

4.1.1. Pyridoxal 5′-phosphate-dependent enzymes.

4.1.1.1. Pyridoxal 5′-phosphate, a versatile enzyme cofactor.

Feeding experiments in rats conducted by György (1893-1976) lead to the identification of vitamin B_6; purification was achieved in 1938.[294] In 1944, Gunsalus, *et al.*[295] discovered that the phosphate ester of the aldehyde form of the vitamin, pyridoxal 5′-phosphate (PLP), could function as an enzyme cofactor. PLP is implicated in > 140 [Ref. [296]] different enzymatic transformations of amino acids, as well as the phosphorylation of glycogen and malodextrin. PLP-dependent enzymes have been exhaustively reviewed.[296-301]

4.1.1.2. PLP metabolism.

All organisms possess enzymes for the supply, trafficking and metabolism of PLP. Phylogenetically ubiquitous is a salvage pathway whereby vitamin B_6, PLP and other vitamers of the B_6 group (pyridoxal, pyridoxamine, PMP) are interconverted. One of the enzymes of the salvage pathway is PdxK, a pyridoxine/pyridoxal kinase that activates vitamin B_6 by phosphorylation.

In addition to the salvage pathway, bacteria, fungi and plants are able to synthesise vitamin B_6 *de novo*. Two synthetic pathways (PDX and SOR) are known; it seems they cannot co-exist in the same organism. Hence, the pyridoxine-requiring (PDX) pathway

is restricted to certain bacteria (e.g., *E. coli*), while the singlet oxygen resistance (SOR) pathway is found in fungi, plants and some bacteria (e.g., *M. tuberculosis*).

The PDX machinery of *E. coli* is well understood and consists of PdxABCFHJ and GapA proteins (reviewed by Spenser and Hill[302]). Conversely, the SOR pathway is inadequately characterised, although two proteins of the pathway are known: Pdx1[303] and Pdx2.[304]

4.1.1.3. PLP metabolism in P. falciparum.

Recent evidence suggests that *de novo* vitamin B_6 synthesis takes place in *P. falciparum* via the SOR pathway. For example, sequence resembling Pdx1 (PlasmoDB MAL6P1.215) was noted in the *P. falciparum* genome by Ralph *et al.*[80] Using high performance liquid chromatography and mass spectrometry, Cassera, *et al.*[104] detected metabolites of the PLP pathway in erythrocytic stage parasites. Extending these findings, Wrenger, *et al.*[282] identified Pdx2 (PlasmoDB chr11.gen_420) and PdxK (GenBank NP_703820) genes in the *P. falciparum* genome while establishing the functionality of malarial Pdx1, Pdx2 and PdxK by complementation tests in yeast and by enzymology with purified protein. None of the SOR genes in *P. falciparum* possess an obvious mitochondrial or plastid-targeting signal, suggesting that vitamin B_6 synthesis may be cytosolic in *P. falciparum*.[282]

4.1.1.4. Chemical structure of PLP.

PLP is composed of a pyridine ring covalently linked to an aldehyde group (CHO) at carbon 1 (C1); a hydroxyl group (OH) at C2; and a phosphate group (PO_4) at C5'.

Evolution has found a use for all these functional groups during PLP-dependent catalysis.

The over-arching chemical feature of PLP-dependent catalysis is substrate oxidation. This process is classically believed to be facilitated by the pyridine ring of PLP, which acts as an electrophile (negative charges readily delocalise through the π-orbitals). In order to draw electrons from the substrate by this mechanism, a covalent bond must be established. This task is fulfilled by the aldehyde group of PLP that forms an imine (Schiff base) bond with the ε-group of the amino acid. Finally, molecular topology must be rigidly maintained during catalysis. This is the function of the OH and PO_4 groups that make hydrogen bonds with residues lining the active site of the enzyme.

4.1.1.5. PLP in enzyme catalysis.

A vast amount of information has accumulated about the mechanism of PLP-dependent catalysis. In part, this is due to the early realisation that certain aspects of PLP catalysis, such as Schiff base formation and electron flux, can be studied non-enzymatically using chemical models.[305] This allowed meaningful experiments to be conducted even before large numbers of PLP-dependent enzymes had been purified.

By convention, the carbon atoms of an amino acid are labelled Cα, Cβ (and in some amino acids, Cγ); it is the Cα that forms a Schiff base with PLP. Hence, it is the action of the PLP electrophile on Cα that is critical for PLP-dependent catalysis.

Withdrawal of electrons from Cα can lead in three major directions.[298] These are: (1) deprotonation of Cα, generating a quinoid intermediate; (2) amputation of the COO^-

group (decarboxylation); or (3) amputation of the group attached to Cβ (aldol cleavage if the group is OH). While many steps in the catalytic mechanism of PLP-dependent enzymes remain mysterious, this initial phase is well understood.

Dunathan[306] postulated that the orientation of the substrate molecule to the PLP molecule controlled the direction of bond breaking (i.e. deprotonation *vs.* decarboxylation *vs.* aldol cleavage). This hypothesis has been confirmed by analysis of X-ray crystal structures of numerous PLP-dependent enzymes.[297]

According to the Dunathan model, the substrate bond under attack must be perpendicular to the plane of the pyridine ring of the PLP. This ensures intimate contact between "frontier" orbitals; the highest energy occupied molecular orbital (HOMO) of the bond under attack interacts with the lowest energy unoccupied molecular orbital (LUMO) of the delocalised π-system of the pyridine ring and Schiff base. Hence, the initial step of the reaction is guided merely by changes in molecular topology, directed by the residues lining the active site of the enzyme.

4.1.1.6. Disparity of fold and function in PLP-dependent enzymes.

Metzler[298] has further classified PLP-dependent enzymes into four main groups (a-d) based on the fate of the substrate. These are: group a (loss of α-hydrogen); group b (decarboxylation); group c (side chain cleavage); and the most complicated, group d (ketimine intermediate as electron acceptor). This mechanistic classification runs in parallel to a classification based on protein fold (see chapter 2). However, the two classifications do not correlate.[297] This obviously has important implications for

bioinformatic analysis: similarity of protein sequence does not imply similarity of function. This explains the need to test the catalytic activity of PfSufS (chapter 5).

4.1.2. Modelling the catalytic mechanism of cysteine desulphurases.

*4.1.2.1. Nomenclature of cysteine desulphurases.**

Cysteine desulphurases (EC 2.8.1.7) are a phylogenetically ubiquitous class of enzymes (chapter 2) that hydrolyse L-cysteine to L-alanine and sulphane sulphur (S^0).[160] They are all PLP-dependent and fall into the aspartate aminotransferase (AAT) category (by similarity of protein sequence) and the d group (by catalytic mechanism). The cellular role of cysteine desulphurases is described in detail in chapter 1; whereas the following discussion is concerned only with structure and mechanism (Fig. 4.1).

4.1.2.2. Molecular structure.

X-ray crystal structures of several prokaryotic cysteine desulphurases have been deposited in the Protein Data Bank (PDB).†

---

* *Note on nomenclature.* As in any rapidly advancing field, nomenclature is confused. For example, some authorities (e.g. Mihara and Esaki[160]) describe cysteine desulphurases as "aminotransferases [sic] class V of fold-type I" while others (e.g. Lima) put them in the "α-family" of PLP-dependent enzymes. For clarity, I have adopted the nomenclature of a recent review.[307] The naming of specific proteins is also difficult. For example, in yeast, IscS is known as Nfs1p.[141] In *E.coli*, SufS is variously known as SufS,[166] CsdB,[308] or even NifS CsdB,[251] while in *Arabiopsis thaliana*, SufS is called NFS2,[198] or CpNifS.[199] Here, I exclusively use the terms IscS and SufS.

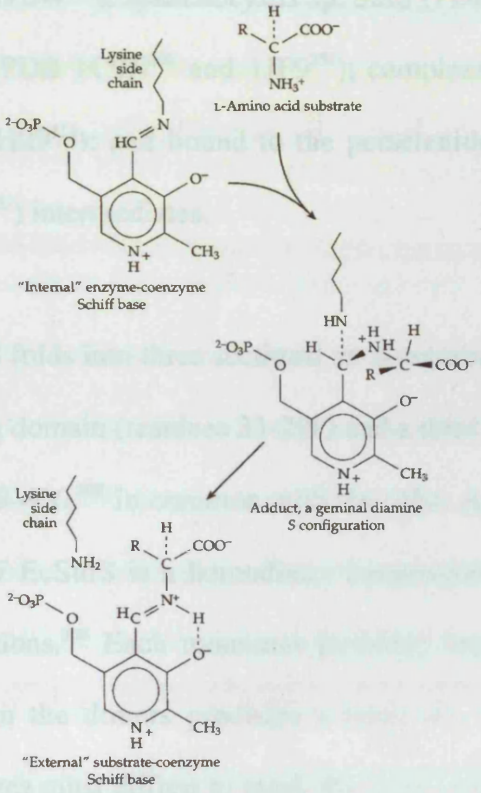
† http://www.rcsb.org/pdb/

FIGURE 4.1.

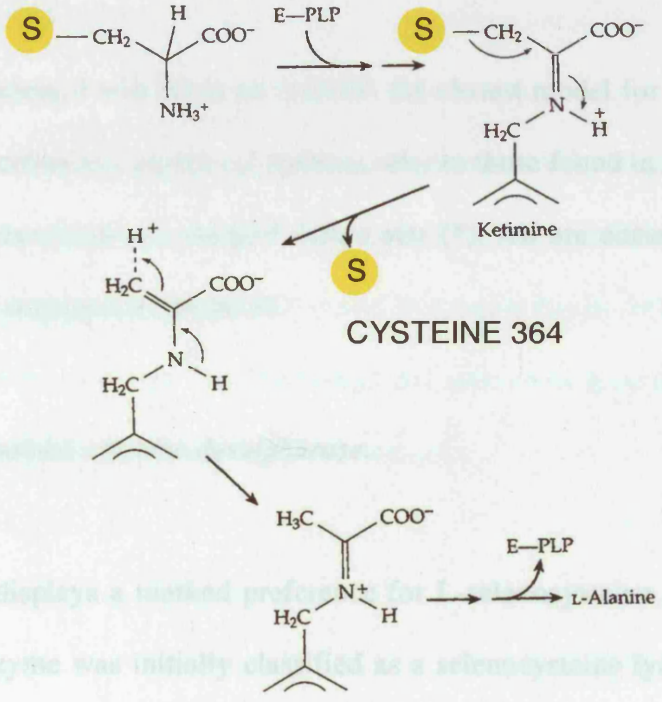
MODEL OF THE CYSTEINE DESULPHURASE REACTION MECHANISM.

Cysteine desulphurase enzymes catalyse the hydrolysis of L-cysteine to sulphur and L-alanine. The process is complex and poorly understood, but can be explained largely by reference to the pyridoxal 5′-phosphate (PLP) cofactor and the active site cysteine of the apoenzyme. A detailed description is given in the text. ***Transimination.*** The bonding environment of the PLP changes, allowing formation of a Schiff base with the L-cysteine substrate. ***γ-elimination via ketimine.*** Electron flux through the substrate drives protonation of the ketimine intermediate. Concurrently, the active site cysteine (not shown) stages nucleophilic attack on the sulphur atom of the substrate (marked in yellow). The sulphur atom is amputated from the substrate and forms a persulphide bond with the active site cysteine. Cleavage of the persulphide bond liberates sulphur.

TRANSIMINATION



γ-ELIMINATION VIA KETIMINE



These include: *Thermatoga maritima* NifS (PDB 1ECX and 1EG5[309]), *E. coli* IscS (PDB 1P3W[310]), *Synechocystis* sp. SufS (PDB 1T3I[311]) and several versions of EcSufS: alone (PDB 1C0N[308] and 1JF9[251]); complexed with an inhibitor (L-propargylglycine) (PDB 1I29[312]); and bound to the perselenide (PDB 1KMK[251]) and persulphide (PDB 1KMJ[251]) intermediates.

EcSufS folds into three sections, an N-terminal segment (residues 1-21), a central PLP-binding domain (residues 33-298) and a third small domain composed of residues 22-32 and 299-406.[308] In common with the other AAT-type enzymes, the catalytically active form of EcSufS is a homodimer constrained by hydrogen bonding and hydrophobic interactions.[308] Each monomer probably binds one molecule of PLP. The interface between the dimers produces a deep (*ca.* 1.5 nm[251]) active site cleft along which substrates must diffuse to reach the active site cysteine and cofactor; histidine residues surrounding the active site probably coordinate the solvent (water).[251] Although residues from one dimer contribute to the active site of the other, the two active sites of the enzyme operate independently.

In the remainder of this review, I will focus on EcSufS, the closest model for PfSufS. Hence, in the account of mechanism, numbered residues refer to those found in EcSufS; residues from the second monomer are marked with a star (*). All are conserved in AtSufS and PfSufS protein sequences (chapter 2).

4.1.2.3. SufS as a complex cysteine desulphurase.

In vitro, isolated EcSufS displays a marked preference for L-selenocysteine over L-cysteine.[313] Indeed, the enzyme was initially classified as a selenocysteine lyase (EC

4.4.1.16).[165] In this context, it is worth noting the selenium starvation experienced by malaria parasites in the red blood cell.[314] As compared to the more active IscS, the lethargic desulphurase activity of SufS against L-cysteine may be ascribed to the shortness and rigidity of the loop surrounding the active site cysteine.

It has recently been shown that accessory factors significantly enhance the cysteine desulphurase activity of SufS. For example, Outten, *et al.*,[191] demonstrated *E. coli* SufE stimulated 8-fold the activity of EcSufS against L-cysteine. Further, addition of SufABC proteins to the SufE-SufS complex augmented the activity 32-fold as compared to SufS alone.[191] Using purified *Erwinia chrysanthemi* SUF proteins expressed in *E. coli*, Loiseau, *et al.*[192] reported that addition of SufE to SufS stimulated cysteine desulphurase activity 50-fold. Hence, *in vivo*, SufS probably operates within a multi-subunit complex.

4.1.2.4. The resting enzyme.

In resting enzymes, PLP is tethered in the active site by a covalent bond (Schiff base) and at least seven hydrogen bonds. The Schiff base linkage is between the carbonyl group of the PLP and the ε-amino group of the catalytic lysine ($K226^{308}$). This bond is discussed in detail in chapter 5. Additionally, a number of active site residues (T95, D200, Q203, S223, H225, K226 and $T278^{*308}$) form hydrogen bonds with PLP. Collectively, these residues maintain the cofactor within the active site after the C=N bond has been hydrolysed to form the enzyme-substrate complex.

4.1.2.5. Formation of the enzyme-substrate complex by transimination.

Transimination, the necessary first step of PLP-dependent reactions, serves to make the cofactor available to the substrate.[298] This involves the hydrolysis of the C=N bond between the cofactor and K226 ("internal" aldimine); and the formation of a new C=N bond between the cofactor and the substrate cysteine ("external" aldimine). In effect, the Schiff base is exchanged. In *Thermatoga maritima* NifS, this leads to the formation of a quinoid intermediate[309] via the Dunathan mechanism.

Transimination can only happen because of the marked lability of the PLP-lysine Schiff base. Indeed, if the "internal" aldimine bond is reduced to a single bond by addition of a metal borohydride, commonly sodium borohydride ($NaBH_4$), the system is locked and the reaction cannot proceed. Hence, $NaBH_4$ serves as a potent irreversible inhibitor of PLP-dependent enzymes and thus is a valuable discriminator for Schiff base participation in reaction mechanisms.[315]

4.1.2.6. γ-elimination via ketimine intermediate.

Cysteine desulphurases catalyse γ-elimination reactions. The process begins with protonation of the Cα of the quinoid intermediate, forming an electrophilic ketimine than can attack Cγ. Polarisation of $C=NH^+$ probably draws electrons through the ketimine intermediate, allowing the stripping of a proton from Cγ by the action of K226.

It is probably at this stage that the active site cysteine (C364) relocates to begin nucleophilic attack on the intermediate,[251]. Briefly, the sulphur atom is confiscated from the substrate, and transferred to C364 where it attaches via a persulphide bond

(persulphide intermediate). However, due to the shortness and inflexibility of the active site loop of SufS, this step may be severely rate-limiting in isolated enzymes.[251] Kinetic evidence from *Synechocystis* sp. SufS suggests that SufE may help the process, either by changing the topology of the active site to facilitate nucleophilic attack by C364; or by attacking the substrate with a nucleophilic cysteine residue of its own.[311]

It is worth noting that the pivotal role of the active site cysteine has been demonstrated in many different cysteine desulphurase enzymes, via mutagenesis,[313] crosslinking,[162] suicide inhibition[316] and alkylation[316] studies. However, mutation of the active site cysteine of *Synechocystis* sp. SufS (C372A) diminished but did not entirely abolish cysteine desulphurase activity.[311] This finding may be attributed to the operation of a second desulphuration pathway in which additional molecules of L-cysteine diffuse into the active site and launch a nucleophilic attack upon the ketimine intermediate independent of the active site cysteine.[311] Action of two distinct cysteine desulphuration pathways *in vitro* could explain the aberrant Michaelis-Menten behaviour observed with EcSufS.[313]

Finally, sulphur is abstracted from C364 onto C51 of SufE, and passed downstream.[191,193] This event is probably helped by R359 of SufS swivelling away from the active site cysteine towards D28, T62* and H55*, thereby improving physical access.[251] The second product, L-alanine, is released; transimination of PLP restores the Schiff base with K226. The active site has completed one catalytic cycle and is ready to process another molecule of L-cysteine.

4.1.2.7. Reversible inhibition by abortive transamination.

The chemical promiscuity of PLP is such that PLP-dependent enzymes often commit unforced mechanistic errors.[301] Hence, in spite of the majority of cysteine desulphurase reactions proceeding in the direction outlined above (formation of sulphur and L-alanine), it seems probable that the reaction will often make a wrong turn.

Abortive transamination is likely to be the predominant fault[313] and probably arises via erroneous protonation at C4′ of the quinoid intermediate (section 4.1.2.6.). Concomitantly, PLP is transformed into inactive pyridoxamine 5′-phosphate (PMP); the cysteine desulphurase becomes reversibly inhibited. Reactivity can be restored by addition of an excess of pyruvate (substrate for reverse transamination, i.e., PMP → PLP) or fresh PLP.[313] Indeed, *in vitro* enzyme assays require addition of free PLP to the buffer in order to sustain the reaction. *In vivo*, abortive transamination probably makes the enzyme exquisitely sensitive to PLP starvation.

4.1.2.8. Kinetic model of the reaction catalysed by SufS.

Although there is a copious general literature on cysteine desulphurases (reviewed by Mihara and Esaki[160]), to my knowledge only four SufS enzymes have been biochemically characterised These are: *E. coli* SufS,[165,313], *Erwinia chrysanthemi* SufS;[192] *Synechocystis* sp. SufS;[311,317] and *A. thaliana* SufS.[198,199]

Such paucity of data has limited the formulation of predictive kinetic models of SufS catalysis. One recent exception is Tirupati, *et al.*,[311] who described a simple model based

on spectroscopic data from *Synechocystis* sp. SufS. The model can be summarised using consecutive reactions constrained by rate constants k_1 and k_2 (Eq. 4.1).[318]

$$A \xrightarrow{k_1} B \xrightarrow{k_2} C \qquad \text{(Eq. 4.1)}$$

Mixing of substrate (*A*) and enzyme leads to formation (k_1) of the persulphide intermediate (*B*). In isolated *Synechocystis* sp. SufS, k_1 is exceptionally slow (0.02 ± 0.005 s^{-1} [Ref. [311]]. It is an amalgam of rapid steps (formation of the enzyme-substrate complex) suddenly braked by highly inefficient persulphide formation (perhaps due to the inflexibility of the active site loop). However, once achieved, the persulphide intermediate breaks down (k_2), releasing sulphur (*C*) and L-alanine; k_2 = 0.0002 $mM^{-1}.s^{-1}$ [Ref. [311]] (second order rate constant). This simple kinetic model has important practical implications for the study of SufS enzymes (section 4.3.2.6.).

4.1.3. Chapter objective and overview.

The objective of the work reported in this chapter was to develop a cysteine desulphurase activity assay suitable for screening heterologously expressed PfSufS (chapter 5). I describe the development of the assay and its validation with *E. coli* SufS and *A. thaliana* SufS enzymes.

4.2. MATERIALS AND METHODS.

4.2.1. Preparation of enzyme controls.

4.2.1.1. Expression and purification of E. coli *SufS.*

E. coli SufS (GenBank P77444) was expressed and purified essentially as described.[191] *E. coli* BL21 codon plus (DE3)-RIL cells (Stratagene) were transformed with the expression plasmid pRSET:His_6-EcSufS (a gift from Dr. F. W. Outten, NIH, USA). Cells harbouring the plasmid were incubated for 16 h in Luria-Bertani (LB) broth (10 g bactotryptone, 5 g yeast extract, 10 g NaCl per 1 L redistilled water) containing 100 $\mu g.mL^{-1}$ carbenicillin. The culture was diluted 1/50 into 1 L of fresh LB broth containing 100 $\mu g.mL^{-1}$ ampicillin. At an OD600 of 0.4, the culture was induced for 3 h at 37°C with 1 mM isopropyl-ß-D-thiogalactopyranoside (IPTG). Cells were harvested by centrifugation and resuspended in 50 mL of extraction buffer containing 50 mM NaH_2PO_4, pH 8.0, 300 mM NaCl, 5 mM imidazole, 1 mM phenylmethylsulphonyl fluoride (PMSF) and Roche complete EDTA-free protease inhibitor cocktail set (1 tablet per 50 mL). The cell suspension was sonicated packed in ice for 1 min in 10 s bursts using a Sonics and Materials Vibracell probe tip sonicator. The lysate was centrifuged at 150000*g* for 30 min at 4 °C and the supernatant was applied to a 1 mL HiTrap HP chelating column (Amersham) charged with Ni^{2+}. After extensive washing, protein was eluted in 250 mM imidazole in extraction buffer. All steps of the purification were done at 4 °C at a flow rate of 1 $mL.min^{-1}$ using an Äkta FPLC (Amersham). Yellow elution fractions containing the enzyme were combined. Any precipitate was re-dissolved by addition of 20 mM dithiothreitol[332] (DTT), and dialysed

for 16 h at 4 °C against 25 mM Tris-HCl, 100 mM NaCl, pH 7.4. Enzyme was stored at –20 °C in 50 % v/v glycerol.

4.2.1.2. Expression and purification of A. thaliana *SufS.*

A. thaliana SufS (GenBank AAL79956) was expressed and purified essentially as described.[198] *E. coli* BL21 cells harbouring the pET20b:AtNFS2-His_6 expression plasmid (a gift from Dr. S. Lobréaux, INRA, Montpellier, France) were incubated for 16 h at 30°C in 1 L LB broth containing 100 $\mu g.mL^{-1}$ ampicillin. D-glucose (1% w/v) was added to suppress expression from the T7 promoter. Cells were harvested by centrifugation and resuspended in 800 mL LB broth containing 50 $\mu g.mL^{-1}$ ampicillin but lacking D-glucose. Protein expression was allowed to proceed for 2 h at 30°C. The bacterial pellet was resuspended in 50 mL of extraction buffer containing 50 mM NaH_2PO_4, 150 mM NaCl, 10 mM imidazole and Roche complete EDTA-free protease inhibitor cocktail set (1 tablet per 50 mL) and sonicated as described (section 4.2.1.1.). The lysate was centrifuged at 10000*g* for 30 min at 4 °C. The supernatant was added to 1 mL of ProBond resin (Invitrogen) previously washed with redistilled water and equilibrated with extraction buffer. The mixture was incubated in a closed tube on a rotator for 1 h at 4 °C. The resin was.packed by gravity into a 10 mL Bio-Rad plastic column prior to washing with 10 mL extraction buffer containing 20 mM imidazole. Protein was eluted in extraction buffer containing 250 mM imidazole. Yellow elution fractions containing the enzyme were stored at –20°C in 50% v/v glycerol.

4.2.1.3. Detection and quantification of expressed proteins.

Expressed proteins were detected by denaturing polyacrylamide gel electrophoresis using the NuPAGE system (Invitrogen) followed by Coomassie brilliant blue staining. The ExPASy ProtParam tool* was used to predict molecular weights and molar extinction coefficients[333] based on amino acid sequence. The predicted molecular weight of both proteins was 48 kDa. Assuming that all the cysteine residues appeared as half cystines, the predicted extinction coefficients at 280 nm were AtSufS-His_6 (37790 $M^{-1}.cm^{-1}$) and His_6-EcSufS (49170 $M^{-1}.cm^{-1}$). Protein concentrations were determined from the absorbance of aromatic amino acids at 280 nm.

4.2.1.4. Assessment of PLP content of expressed proteins.

The PLP content of expressed proteins was assessed by inspection of absorbance spectra for a characteristic Schiff base peak in the yellow wavelengths. Additionally, the response of the PLP cofactor to L-cysteine was determined. Using a Beckman DU 640 spectrophotometer, a spectral scan (300-600 nm) was acquired of 98 μL of enzyme (2 $mg.mL^{-1}$) contained in a quartz-glass cuvette. A small volume (2 μL) of 25 mM L-cysteine (Sigma) was added to the cuvette and the mixture stirred with a pipette tip, giving a final L-cysteine concentration of 0.5 mM. The reaction was monitored over time by acquisition of spectra.

* http://www.expasy.org/tools/protparam.html

4.2.2. Cysteine desulphurase activity by lead sulphide zymogram assay.

Cysteine desulphurase activity of purified *E. coli* IscS (a gift from Dr. S. Adinolfi, NIMR) was assessed using the lead sulphide zymogram method essentially as described by Borup.[321] Proteins were separated by denaturing polyacrylamide gel electrophoresis using the NuPAGE system. Critically, protein was left unheated prior to loading on the gel. After the gel had been run, the enzyme was re-natured by removal of sodium dodecyl sulphate (SDS). This was achieved by washing gels twice in redistilled water and twice in 50 mM Tris-HCl, pH 7.4; both washing buffers contained 2.5% v/v Triton X-100. Finally, gels were washed twice in 50 mM Tris-HCl, pH 7.4 without detergent and incubated at 37 °C for 16 h in freshly-made developing solution containing 50 mM Tris, pH 7.4, 10 mM $MgCl_2$, 10 mM L-cysteine, 0.5 mM lead nitrate ($PbNO_3$), 5 mM DTT and 5 mM PLP. Active bands were identified by the presence of precipitated lead sulphide (PbS).

4.2.3. Cysteine desulphurase activity by methylene blue assay.

4.2.3.1. Enzyme assay.

Reactions contained 50 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 10 μM PLP, 100 μM L-cysteine, 70-1000 nM of enzyme and 5 mM of reductant. Reductants used were: ß-mercaptoethanol (ß-ME); DTT; tris(2-carboxyethyl)phosphine[334] (TCEP) (Novagen); and tris(hydroxypropyl)phosphine (THP) (Calbiochem). Incubation was for 2-15 h in 1 x 10 cm Pyrex glass test tubes on the open bench (*ca.* 20 °C). Reactions were stirred with a magnetic stirring flea and incubated under 1 cm depth of mineral oil (Sigma) to

reduce the loss of H_2S (g) expected after addition of the methylene blue assay reagents.[335]

4.2.3.2. Determination of sulphide concentration by methylene blue assay.

Sulphide output was measured according to the endpoint methylene blue method of Siegel.[326] Equal volumes (100 μL) of 20 mM N,N-dimethyl-*p*-phenylenediamine sulphate (DMPD) in 7.2 M HCl and of 30 mM $FeCl_3$ in 1.2 M HCl were injected under the oil layer. The mixture was immediately stirred using the magnetic stirring flea. Escape of H_2S (g) was monitored by the "bad egg" odour. After a brick-red flush, a blue colour (methylene blue) rapidly (< 30 min) developed, indicating the presence of sulphide. The mixture was centrifuged at 16000*g* to remove precipitated protein. Absorbance spectra were obtained on a Beckman DU 640 spectrophotometer. Absorbance at 620 nm was obtained on a Dynex Technologies MRX TC II plate reader.

4.2.3.3. Recovery of sulphide by zinc acetate precipitation.

Sulphide was recovered using the method of Gilboa-Garber.[336] Briefly, aqueous solutions of 2.6 % w/v zinc acetate and 6 % w/v NaOH were rapidly mixed in a ratio of 5:1. This mixture (1 mL) was immediately added to the enzyme assay solution. The tube was sealed with a rubber septa stopper, shaken and centrifuged at 3000*g* for 10 min. The resulting pellet (zinc sulphide) was washed with 5 mL of 1.5 M NaCl, followed by 5 mL of redistilled water. Both solutions had been adjusted to pH 8 with NaOH immediately prior to use. Finally, the pellet was resuspended in redistilled water and submitted to the methylene blue assay (section 4.2.3.2).

4.2.3.4. Preparation of sodium sulphide standard solution.

Sodium sulphide (Na_2S) standard solution was prepared immediately before use. To remove residues of sodium sulphite (Na_2SO_3), crystals of Na_2S (Aldrich) were briefly washed with redistilled water (adjusted to pH 8 with NaOH). A stock solution of Na_2S was made to a concentration of 100 mM using redistilled water adjusted to pH 8 with NaOH. Escape of H_2S (g) was monitored by the "bad egg" odour.

4.3. RESULTS AND DISCUSSION.

4.3.1. Preparation of enzyme controls.

4.3.1.1. Overexpression and purification of enzymes.

His_6-tagged EcSufS and AtSufS were overexpressed in *E. coli.* Cells overexpressing EcSufS grew slowly before induction (Fig. 4.2, inset), presumably due to unregulated expression of active enzyme exhausting the supply of L-cysteine. EcSufS-expressing cell pellets were bright yellow because of the accumulation of PLP.[162]

Total soluble protein was extracted from the cells and His_6-tagged protein was purified over Ni^{2+} resin using a simple one-step elution. A typical 280 nm elution trace for EcSufS is presented in Fig. 4.2A. A Coomassie-stained polyacrylamide gel of AtSufS purification fractions is presented in Fig. 4.3A.

Following the binding step of EcSufS purification, a bright yellow band was visible in the Ni^{2+} resin. The yellow band persisted through washing but disappeared from the column on elution. The band probably represented enzyme-bound PLP. Elution of EcSufS with 250 mM imidazole immediately caused protein precipitation. Further, the yellow colour of the solution diminished in brightness probably due to oxidation of the PLP. Eluted protein was re-dissolved (and yellow restored) by reduction with 20 mM DTT. It is interesting to note that another reducing agent (THP) was unable to restore the yellow colour (section 4.3.2.6.).

FIGURE 4.2.

PURIFICATION AND PROPERTIES OF *ESCHERICHIA COLI* SUFS.

A. Ultraviolet absorbance (280 nm) profile showing elution of *E. coli* SufS.

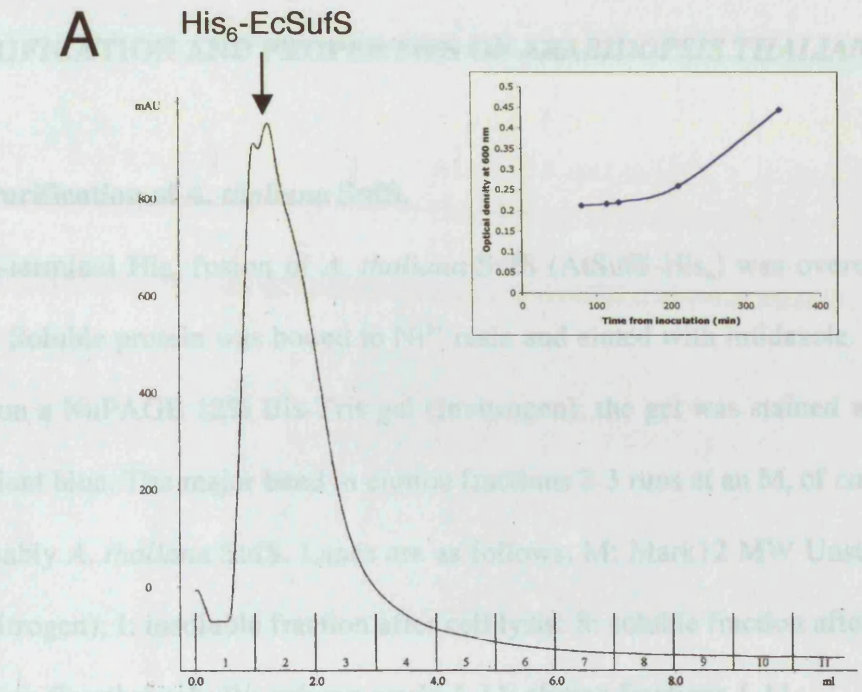
An N-terminal His_6 fusion of *E. coli* SufS (EcSufS-His_6) was overexpressed in bacteria. Soluble protein was bound to Ni^{2+} resin and eluted with imidazole using an Äkta FPLC (Amersham). ***Inset.*** Pre-induction growth curve of bacteria overexpressing *E. coli* SufS. An overnight culture was diluted 1/50 into fresh medium and grown at 37 °C. Uninduced cells took *ca.* 6 h incubation to reach an OD600 of 0.4 (induction point).

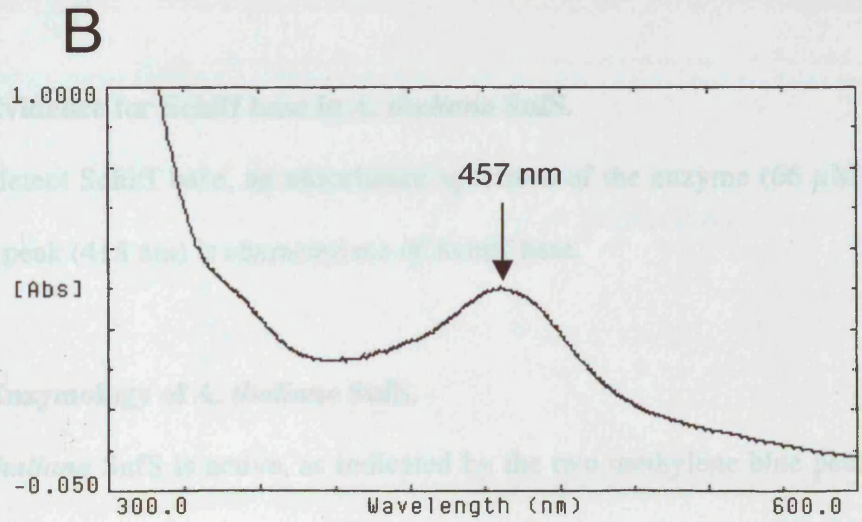
B. Absorbance spectrum of *E. coli* SufS in the yellow regime of the spectrum.

To detect the Schiff base feature characteristic of PLP-dependent enzymes, an absorbance spectrum in the yellow regime was acquired for *E. coli* SufS (70 μM).

C. Enzymology of *E. coli* SufS.

E. coli SufS (700 nM) was shown to be enzymatically active using the methylene blue assay. Dual absorbance peaks at 667 nm and 744 nm are diagnostic of methylene blue.


A
His6-EcSufS
mAU
800
600
400
200
0
0.0
2.0
4.0
6.0
8.0
ml
1
2
3
4
5
6
7
8
9
10
11
Optical density at 600 nm
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0
100
200
300
400
Time from inoculation (min)


B
1.0000
457 nm
[Abs]
-0.050
300.0
Wavelength (nm)
600.0

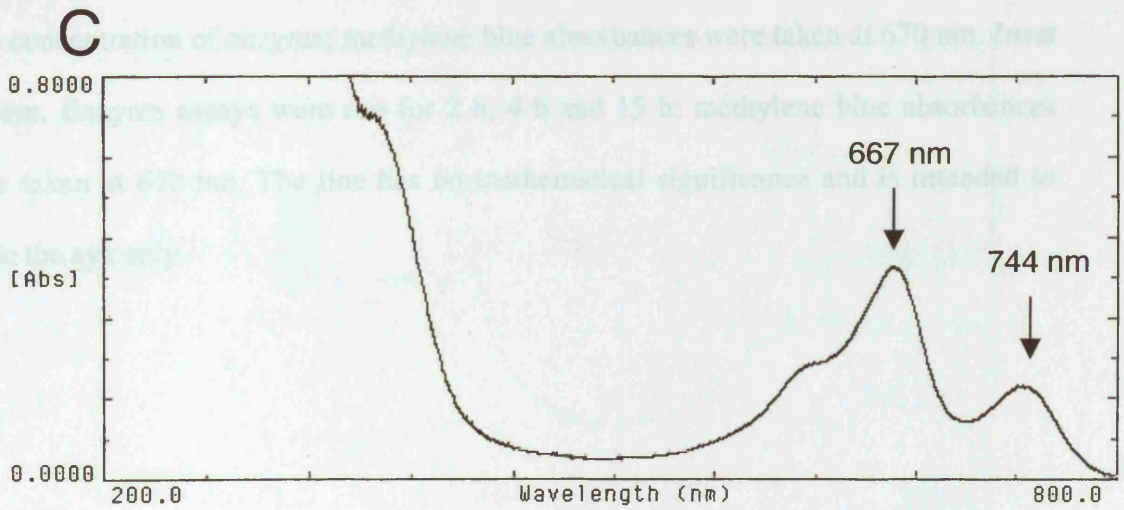

C
0.8000
667 nm
744 nm
[Abs]
0.0000
200.0
Wavelength (nm)
800.0

FIGURE 4.3.

PURIFICATION AND PROPERTIES OF *ARABIDOPSIS THALIANA* SUFS.

A. Purification of *A. thaliana* SufS.

A C-terminal His_6 fusion of *A. thaliana* SufS (AtSufS-His_6) was overexpressed in *E. coli.* Soluble protein was bound to Ni^{2+} resin and eluted with imidazole. Fractions were run on a NuPAGE 12% Bis-Tris gel (Invitrogen); the gel was stained with Coomassie brilliant blue. The major band in elution fractions 2-3 runs at an M_r of *ca.* 48 kDa and is probably *A. thaliana* SufS. Lanes are as follows. M: Mark12 MW Unstained Standard (Invitrogen); I: insoluble fraction after cell lysis; S: soluble fraction after cell lysis; FT: column flowthrough; W: column wash; 1-11: elution fractions 1-11.

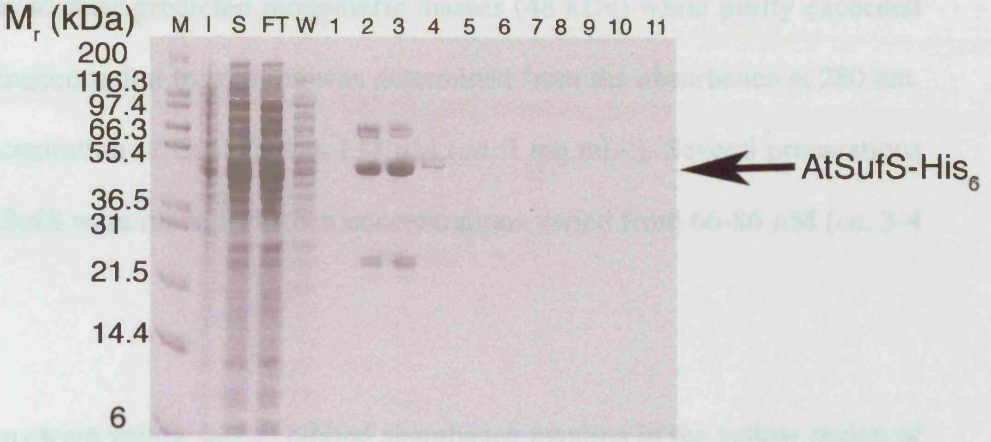
B. Evidence for Schiff base in *A. thaliana* SufS.

To detect Schiff base, an absorbance spectrum of the enzyme (66 μM) was acquired. The peak (418 nm) is characteristic of Schiff base.

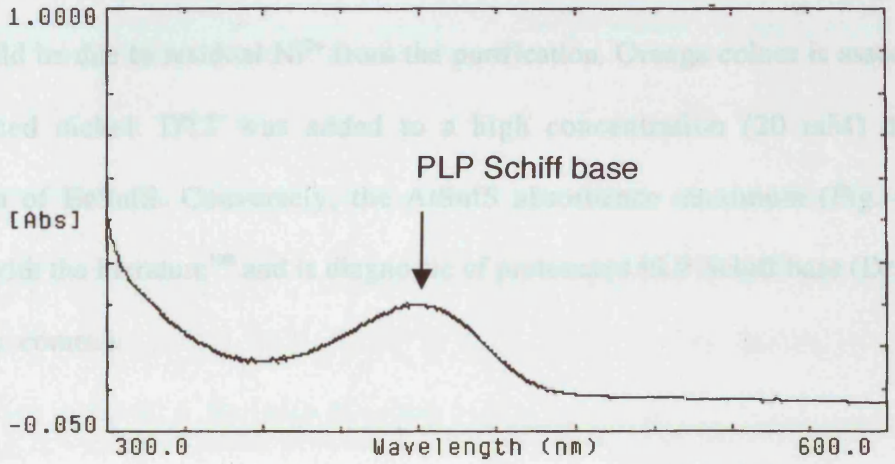
C. Enzymology of *A. thaliana* SufS.

A. thaliana SufS is active, as indicated by the two methylene blue peaks (667 nm and 744 nm). ***Inset top.*** Enzyme assays were performed with a low (70 nM) and high (700 nM) concentration of enzyme; methylene blue absorbances were taken at 670 nm. ***Inset bottom.*** Enzyme assays were run for 2 h, 4 h and 15 h; methylene blue absorbances were taken at 670 nm. The line has no mathematical significance and is intended to guide the eye only.

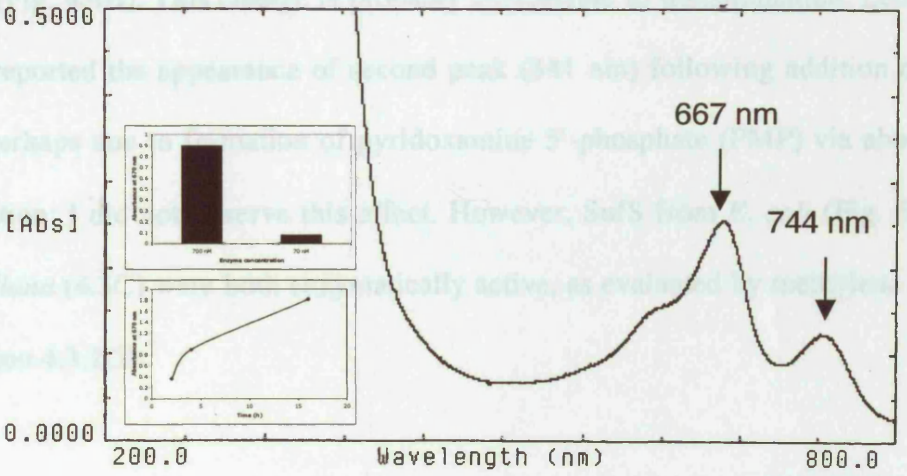
A


M r (kDa)
M I S FT W 1 2 3 4 5 6 7 8 9 10 11
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
AtSufS-His6

B


1.0000
[Abs]
-0.050
PLP Schiff base
300.0
Wavelength (nm)
600.0

C


0.5000
[Abs]
0.0000
667 nm
744 nm
200.0
Wavelength (nm)
800.0

4.3.1.2. Characterisation of purified protein.

On a Coomassie-stained acrylamide gel (Fig. 4.5A), both SufS proteins migrated with mobilities close to their predicted monomeric masses (48 kDa) while purity exceeded 95%. Protein concentration in solution was determined from the absorbance at 280 nm. The stock concentration of EcSufS was 137 μM (*ca.* 7 mg.mL^{-1}). Several preparations of *A. thaliana* SufS were made and stock concentrations varied from 66-86 μM (*ca.* 3-4 mg.mL^{-1})

Protein solutions were yellow and exhibited absorbance maxima in the yellow region of the spectrum (457 nm for EcSufS; 418 nm for AtSufS). The significantly red-shifted peak of EcSufS (Fig. 4.2B), as compared to the reported value for the enzyme (420 nm[165]), could be due to residual Ni^{2+} from the purification. Orange colour is associated with reduced nickel: DTT was added to a high concentration (20 mM) during preparation of EcSufS. Conversely, the AtSufS absorbance maximum (Fig. 4.3B) coincides with the literature[198] and is diagnostic of protonated PLP Schiff base (Dr. A. J. Iriarte, pers. comm.).

Mixing of L-cysteine with AtSufS produced a characteristic[198] red shift of the Schiff base peak (Fig. 4.4B). This change is probably attributable to transimination. Léon, *et al.*[198] also reported the appearance of second peak (341 nm) following addition of L-cysteine, perhaps due to formation of pyridoxamine 5′-phosphate (PMP) via abortive transamination; I did not observe this effect. However, SufS from *E. coli* (Fig. 4.2C) and *A. thaliana* (4.3C) were both enzymatically active, as evaluated by methylene blue assay (section 4.3.2.3).

FIGURE 4.4.

EXPERIMENTS WITH PURIFIED *ARABIDOPSIS THALIANA* SUFS.

A. Purified *E. coli* SufS and *A. thaliana* SufS.

Purified *E. coli* SufS and *A. thaliana* SufS proteins were run on a NuPAGE 12% Bis-Tris gel (Invitrogen); the gel was stained with Coomassie brilliant blue. Lanes are as follows. M: Mark12 MW Unstained Standard (Invitrogen); Ec: *E. coli* SufS; At: *A. thaliana* SufS.

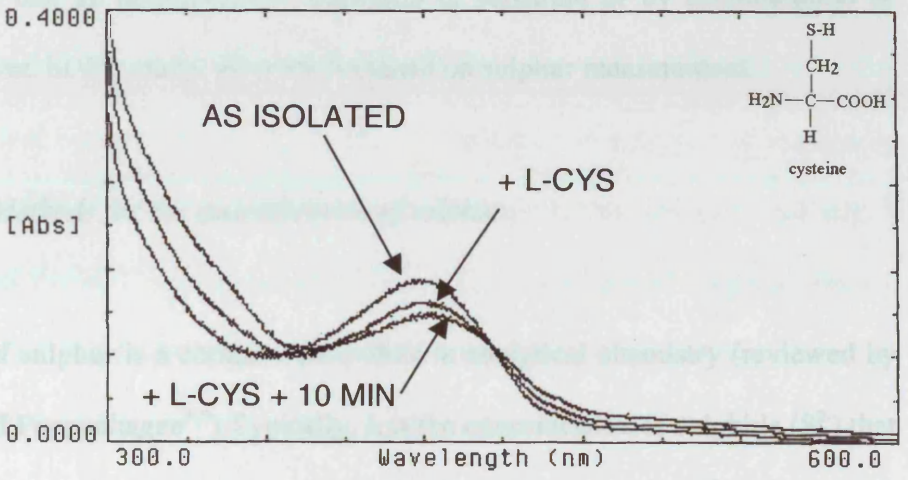
B. Effect of L-cysteine on the PLP Schiff base of *A. thaliana* SufS.

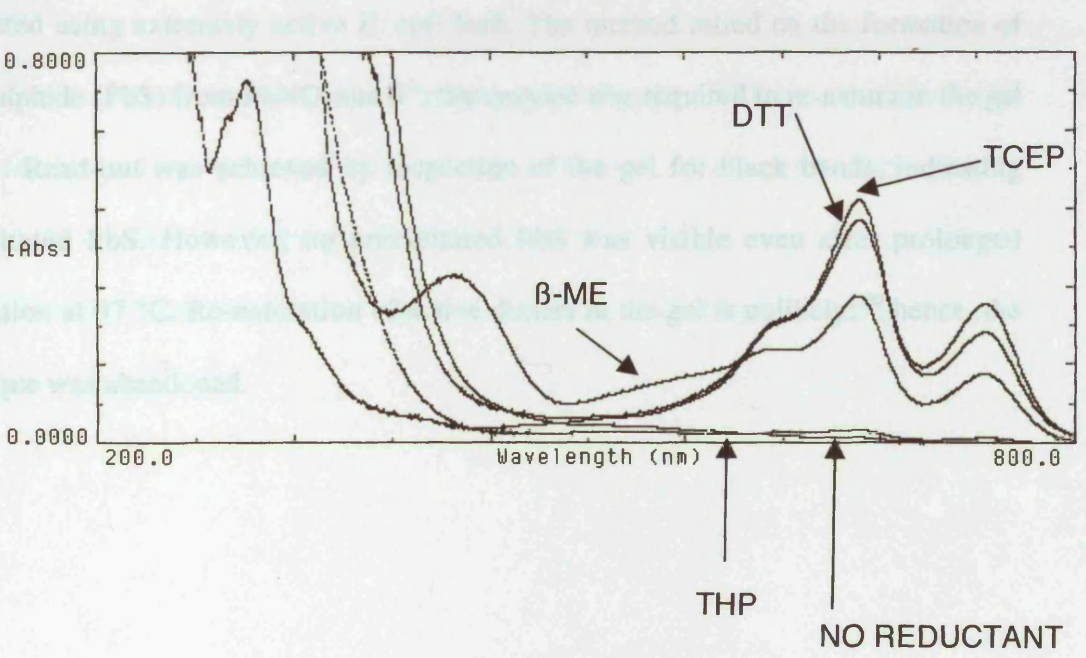
The response of the enzyme (2 $mg.mL^{-1}$) to L-cysteine was determined as described.[198] L-cysteine (0.5 mM) was added and spectra taken at the outset (+ L-CYS) and 10 min later (+ L-CYS + 10 MIN). ***Inset.*** Molecule of cysteine.

C. Effect of reducing co-substrates on the enzymatic activity of *A. thaliana* SufS.

Enzyme assays were set up with *A. thaliana* SufS (900 nM) and 5 mM of one of the following reductants: ß-ME, DTT, TCEP or THP. After 3 h incubation, enzymatic reactions were subjected to the methylene blue test for sulphide.


A
M$_r$ (kDa)
M Ec At
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
SufS


B
0.4000
0.0000
[Abs]
AS ISOLATED
+ L-CYS
+ L-CYS + 10 MIN
S-H
CH2
H2N—C—COOH
H
cysteine
300.0
Wavelength (nm)
600.0


C
0.8000
0.0000
[Abs]
DTT
TCEP
ß-ME
200.0
Wavelength (nm)
800.0
THP
NO REDUCTANT

During storage, activity was preserved by mixing the enzyme solution 50 % v/v with glycerol and freezing at –20 °C. (Activity was lost if the enzyme was frozen without glycerol.)

4.3.2. Development of the cysteine desulphurase assay.

4.3.2.1. Detecting cysteine desulphurase activity.

Given that cysteine desulphurases hydrolyse L-cysteine to L-alanine and sulphur, enzyme activity can be monitored by depletion of substrate or by accumulation of products. However, in this study, all work focussed on sulphur measurement.

4.3.2.2. Methods for the quantification of sulphur.

Quantification of sulphur is a common procedure in analytical chemistry (reviewed by Patterson Jnr and Pappenhagen[319]) Typically, it is the concentration of sulphide (S^{2-}) that is measured.[320] In this study, a lead sulphide (PbS) zymogram method[321] was initially attempted using extremely active *E. coli* IscS. The method relied on the formation of lead sulphide (PbS) from $PbNO_3$ and S^{2-}; the enzyme was required to re-nature in the gel matrix. Read-out was achieved by inspection of the gel for black bands, indicating precipitated PbS. However, no precipitated PbS was visible even after prolonged incubation at 37 °C. Re-naturation of active dimers in the gel is unlikely;[322] hence, the technique was abandoned.

4.3.2.3. Methylene blue method.

A second approach was adopted, the methylene blue (MB^+) method. First reported by Fischer (1852-1919), the MB^+ reaction entails the dimerization of N,N-dimethyl-*p*-phenylenediamine (DMPD) by the action of sulphide and oxidising ferric ions (Fe^{3+}), to form methylene blue, a heterocyclic thiazine dye.[320] The assay is probably the most sensitive colorimetric method for the detection of sulphide.[319] However, the chemistry of the reaction is poorly understood, although mechanisms have been proposed (Fig. 4.5A).[323,324]

It is interesting to note that modern versions of the MB^+ assay employ anodic oxidation in place of Fe^{3+} and thus are able to relate the concentration of sulphide to changes in electrical current.[325] Nevertheless, colorimetric detection remains adequate. Initially, I demonstrated that the MB^+ signal was proportional to the amount of sulphide (Beer's law), using Na_2S as a standard (Fig. 4.5B). It is important to note that the scatter of points in Fig. 4.5B indicates the escape of H_2S (g).

The MB^+ technique was adapted for biochemistry by Siegel[326] and used for the measurement of cysteine desulphurase activity by Zheng, *et al.*[162] The assay operates in three stages: (1) enzymatic conversion of L-cysteine to L-alanine and sulphur by the action of the cysteine desulphurase; (2) reduction of S^0 to S^{2-} by a reducing co-substrate, e.g., DTT; and (3) formation of MB^+ following addition of developers (DMPD, Fe^{3+} and HCl).

FIGURE 4.5.

PHYSICAL CHEMISTRY OF THE METHYLENE BLUE TEST FOR SULPHIDE.

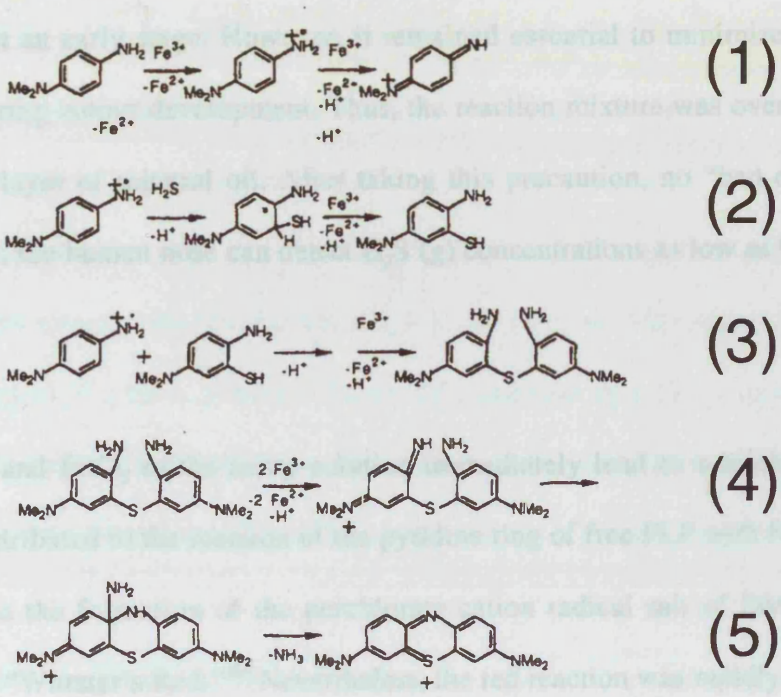
A. Model of the formation of methylene blue from N,N-dimethyl-*p*-phenylenediamine by the action of sulphide and ferric ions.

Kuban, *et al.*[324] envisaged a six-electron oxidation over five steps (1-5). The reaction is started by the rapid mixing of sulphide (S^{2-}) with N,N-dimethyl-*p*-phenylenediamine (DMPD), ferric ions (Fe^{3+}), and an aggressive proton donor (HCl). DMPD is oxidised by Fe^{3+} (1). S^{2-} enters (2) in the form of H_2S, generated by the action of HCl. The cation radical dimerizes (3) and is oxidised by Fe^{3+} (4). Finally, ammonia is eliminated, giving methylene blue (5).

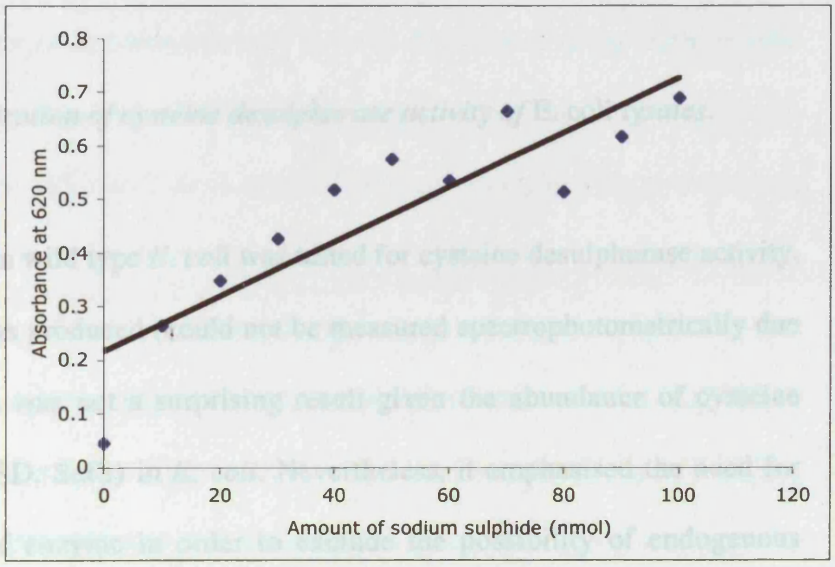
B. Methylene blue production obeys Beer's law.

The production of methylene blue is proportional to the amount of sulphide (range 10-100 nmol). Scatter is due to escape of H_2S (g). Neutral solutions of Na_2S were prepared and immediately subjected to the methylene blue test, as described (section 4.2.3.2). Absorbances were read at 620 nm. The line is intended to guide the eye.

A



B



Amount of Na_2S (nmol) 0 10 20 30 40 50 60 70 80 90 100

Enzyme assays were commenced using *E. coli* and *A. thaliana* SufS. In my hands, the enzymatic reaction did not appear oxygen sensitive, as is commonly assumed (Dr. S. Adinolfi, pers. comm.). Indeed, scrupulous degassing and flushing of solutions with N_2 (g) was abandoned at an early stage. However, it remained essential to minimise the escape of H_2S (g) during colour development. Thus, the reaction mixture was overlaid with a thick (1 cm) layer of mineral oil. After taking this precaution, no "bad egg" odour was detectable; the human nose can detect H_2S (g) concentrations as low as 0.02 parts per million.[327]

Addition of DMPD and $FeCl_3$ to the assay solution immediately lead to a brick-red flush. This may be attributed to the reaction of the pyridine ring of free PLP with $FeCl_3$ ("Phenol test"); or to the formation of the perchlorate cation radical salt of DMPD, commonly known as "Wurster's Red."[324] Nevertheless, the red reaction was rapidly out-competed by dimerization of the DMPD cation radical to form methylene blue.

4.3.2.4. Demonstration of cysteine desulphurase activity of E. coli *lysates.*

Total cellular lysate from wild type *E. coli* was tested for cysteine desulphurase activity. A strong MB^+ signal was produced (could not be measured spectrophotometrically due to high turbidity). This was not a surprising result given the abundance of cysteine desulphurases (IscS, CSD, SufS) in *E. coli*. Nevertheless, it emphasised the need for rigorous purification of enzyme in order to exclude the possibility of endogenous enzyme activity. To confirm that sulphide was absent in prepared cellular lysates and that this signal was a genuine result of enzymatic action, an aliquot of lysate was inactivated before use in the assay by heating to 70 °C for 10 min. Treated material did not produce an MB^+ signal, indicating that the preparation lacked sulphide.

4.3.2.5. Assessment of assay sensitivity.

The sensitivity of the enzyme assay was examined. Assays were performed with a low (70 nM) and high (700 nM) concentration of AtSufS; MB^+ absorbances were taken at 670 nm. The higher concentration of enzyme produced a strong blue signal while the lower concentration produced a signal near background (Fig 4.3C, inset top). These data suggest that enzyme concentrations ~ 1 μM are required for unambiguous detection of SufS activity. Enzyme assays were also run for 2 h, 4 h and 15 h and MB^+ absorbances read at 670 nm (Fig. 4.3C, inset bottom). These data indicate that the L-cysteine substrate is not limiting over short periods (< 4 h) while longer incubation times (15 h) increase the sensitivity of the assay.

4.3.2.6. Evaluation of reducing co-substrates.

Reductants can enhance cysteine desulphurase activity. For example, experiments with *E. coli* CSD showed that inclusion of DTT in the enzyme assay raised the k_{cat} for the release of sulphur from 0.5 min^{-1} to 3 min^{-1}.[328] Recent experiments conducted by Behshad, *et al.*[329] with *Synechocystis* sp. IscS showed that the organophosphine reductant TCEP further enhanced cysteine desulphurase activity (second order rate constant = 4.9 x 10^3 $M^{-1}.s^{-1}$), as compared to the thiolates ß-mercaptoethanol (2.0 $M^{-1}.s^{-1}$) and DTT (340 $M^{-1}.s^{-1}$). This discovery was attributed to the greater capacity of TCEP to cleave the persulphide intermediate within the enzyme active site.

Therefore, I evaluated different reducing co-substrates for their ability to promote the activity of AtSufS. Critically, the enzyme had been isolated without exposure to reducing agents. The organophosphine THP was tested in addition to the reductants

used by Behshad, *et al.*[329] It is worth noting that TCEP and THP powerfully cleave disulphide bonds: TCEP has a reduction capacity against oxidised glutathione of 0.182 milliequivalents.g^{-1}; the figure for THP is 0.232 milliequivalents.g^{-1}.*

The results are reported in Fig. 4.4C. MB^+ signal was detectable in the TCEP, DTT and ß-mercaptoethanol treatments (signal strength: TCEP > DTT >> ß-mercaptoethanol). ß-mercaptoethanol gave a persistent brick-red tinge to the MB^+ solution (probably represented by the pronounced peak in the violet regime of the absorbance spectrum). The colour was likely due to a competing reaction between the mercaptan and DMPD. It is well known that sulphur-containing compounds (e.g., thiols) can interfere with the MB^+ reaction.[319] According to Kuban, *et al.*,[324] the side reaction produces a sulphide-substituted cation radical possessing a chromophore. In order to bypass this interference effect, attempts were made to recover the sulphide by zinc acetate precipitation. However, extra manipulations inevitably lead to loss of sulphide; indeed, no MB^+ signal was detectable following the procedure.

TCEP (Abs 667 nm = 0.50) appeared to offer only a minor enhancement of signal strength over DTT (Abs 667 nm = 0.46). This result apparently conflicts with the finding of Behshad, *et al.*[329] for IscS. However, my result can be readily explained by reference to the kinetic model (Eq. 4.1): in isolated SufS, the rate limiting step is nucleophilic attack (k_1) upon the substrate; while in IscS, the rate limiting step is cleavage of the persulphide intermediate (k_2). Hence, acceleration of k_2 by use of TCEP could only have a minor impact on the activity of AtSufS because of the bottleneck imposed by k_1.

* Calbiochem Data Sheet 598250 Rev. 22-January-04.

In spite of its significant reducing power, THP did not produce an MB^+ signal. One speculation is that THP is unable to cleave the persulphide intermediate due to its flexible side chains impairing entry to the active site. It is worth noting that THP was incapable of reducing the PLP cofactor resident in the active site of resting EcSufS (section 4.3.1.1).

4.3.3. Future work.

In conclusion, I have developed a robust cysteine desulphurase assay based on the methylene blue test for sulphide. Availability of active *P. falciparum* SufS would allow classical enzymological studies to be conducted, e.g. determination of maximum velocity (V_{max}) and Michaelis constant (K_m). Availability of other members of the SUF complex from *P. falciparum* (e.g. SufE) would allow stimulation experiments. However, Mihara, *et al.*[313] were unable to determine the kinetic parameters of EcSufS because enzyme activity against L-cysteine departed markedly from Michaelis-Menten kinetics (determined by inspection of a Lineweaver-Burk plot).

Given the number of PLP-dependent enzymes, there are numerous examples of inhibitors, although the majority are non-specific (reviewed by Eliot and Kirsch[297]). Many are potent poisons forming stable complexes with PLP (e.g., hydrazine). A few have been developed for clinical use, e.g., eflornithine, which attacks ornithine decarboxylase (EC 4.1.1.17) and is used to treat African trypanosomiasis caused by *Trypanosoma brucei gambiense*.[330] Hence, a screen for inhibitors of *P. falciparum* SufS might be useful. Two obvious candidates would be L-allylglycine and L-vinylglycine, suicide inhibitors already used to probe the cysteine desulphuration mechanism of *Azotobacter vinelandii* NifS.[316] If a specific inhibitor were identified, a chemical

knockout of SufS in living parasites could be attempted, as demonstrated with falcipain-1.[331]

CHAPTER 5.

EXPRESSION AND PURIFICATION OF RECOMBINANT *PLASMODIUM FALCIPARUM* SUFS.

P. falciparum SufS (PfSufS) is predicted to be a cysteine desulphurase. In order to test this prediction, attempts were made to over-express recombinant PfSufS in heterologous hosts. This chapter briefly summarises the problems encountered with this technique, followed by a description of the numerous experiments conducted to obtain pure and active PfSufS.

5.1. INTRODUCTION.

5.1.1. Studying enzyme activity.

In enzymology, it is vital to use an homogenous preparation so that activity can be unequivocally attributed to a single protein. The purpose of this brief introduction is to outline how *P. falciparum* enzymes may be obtained.

5.1.2. Isolation of authentic enzyme.

The classical method is to isolate authentic enzyme directly from the organism under study. Such an approach requires careful attention to the physical properties of the protein, notably, size, polarity and solubility, to allow purification from a complex mixture. Amenable to this technique are proteins expressed at high levels; or for which a large amount of starting material is available.[337] Application to malarial enzymes was facilitated by the development of *in vitro* culture of *P. falciparum* erythrocytic stages.[338] For example, Chen and Zolg[339] purified microgram quantities of *P. falciparum* dihydrofolate reductase-thymidylate kinase (PfDHFR-TS; PlasmoDB PFD0830w) through eight steps from a large number ($\sim 1.5 \times 10^{11}$) of erythrocytic stage parasites. However, given the time and effort required, such techniques have fallen out of favour as compared to recombinant expression (section 5.1.3). To my knowledge, there are no reports of successful purification of plastid-targeted enzymes from *P. falciparum* lysates. It seems probable that the low levels of expression associated with these enzymes (chapter 1) and their imprisonment by multiple membranes (chapter 3) would make successful recovery challenging.

5.1.3. Recombinant expression.

5.1.3.1. Over-expression systems.

An alternative approach is to clone the gene of interest into a vector along with an inducible promoter and use the resulting plasmid to programme the over-expression of the protein in large quantities, *in vitro*, or in prokaryotic[340] or eukaryotic hosts.[341] The bulk of this work done in *E. coli* produces enzymes that are usually considered good models for authentic enzyme. For example, authentic and recombinant versions of PfDHFR-TS (section 5.1.2) exhibited similar kinetic properties.[342] Likewise, over-expression in *E. coli* has been used successfully to obtain active SufS from diverse organisms.[165,192,198,199,311]

The major advantage of the heterologous approach becomes apparent during purification. As the protein is already heavily enriched by over-expression, it is easier to separate from a complex mixture. Furthermore, genetic fusion to an affinity tag helps purification. For example, poly-histidine tags bind to divalent metal cations (e.g., Ni^{2+}), enabling immobilised metal ion chromatography. Indeed, this approach can allow purification of a protein of interest to homogeneity in one step (e.g., His_6-EcSufS, chapter 4).

Nevertheless, the heterologous approach often fails with malarial proteins, due to problems of transcription, translation and protein folding arising from incompatibilities between *P. falciparum* proteins and those of the expression host. It is illustrative to note the small number (~ 20-30) of *P. falciparum* proteins that have been successfully over-expressed in the last 20 y.[343] Such anecdotal evidence is supported by high throughput

screening, e.g., only 149 (22 %) of 672 *P. falciparum* proteins made by the Structural Genomics of Parasitic Protozoa (SGPP*) consortium are reported to be soluble (15 April 2005).

5.1.3.2. Problems encountered at the level of expression.

P. falciparum genes are characterised by a high AT content (80.6 %),[45] markedly divergent from *E. coli* (49.2 % AT) and eukaryotic expression hosts (e.g., *S. cerevisiae*, 61.7 % AT). Hence, expression often fails due to a lack of appropriate tRNAs. In *E. coli*, the problem can be overcome to some extent by co-transformation with helper plasmids encoding the required tRNAs, e.g., the RIG plasmid.[344] In the case of eukaryotic expression systems, it is necessary to re-synthesise the *P. falciparum* gene according to the codon usage of the expression host.[345] Even in *E. coli*, re-synthesis can offer significant improvements to the expression level, e.g., the ten-fold efficiency gain in expression of synthetic PfDHFR-TS over wild-type.[346]

5.1.3.3. Problems encountered at the post-translational level.

If protein is expressed, problems can still arise. Notably, the protein may be toxic to the host; or it may adhere to membranes, requiring extraction with detergents. However, the most widely publicised obstacle is protein insolubility typified by the formation of inclusion bodies (IB) in the cytosol or periplasm.[347]

IB may be crudely defined as a dense protein aggregate that partitions into the pellet following cellular lysis and high-speed centrifugation (e.g., 150000*g*). However, it is

* http://depts.washington.edu/sgpp/

worth noting that lysis and centrifugation conditions have a marked influence on this partitioning; a more exact definition may be derived from recent investigations.[348] IB was found to be an aggregate of disordered structures and ß-sheets, reminiscent of amyloid plaques of eukaryotic cells.

Certain physical properties of proteins are related to aggregation, e.g., cysteine fraction, net charge, hydrophilicity and propensity to form turns. For eukaryotic proteins, the ultimate cause of aggregation may be incompatibility between eukaryotic protein folding pathways and the alien environment of *E. coli*. For example, transcription and translation are coupled in bacteria, and the cytosol of *E. coli* is basic (pH 7.5-7.9), reducing, and lacking in eukaryotic chaperones.[349]

Remaining sections of this introduction focus on overcoming the IB problem *in vivo* and *ex vivo*. Given that a general method to ensure correct protein folding is not available, methodology is empirical[350]

5.1.4. Protein solubilisation *in vivo*.

5.1.4.1. Solubilisation by manipulation of expression conditions.

If a protein forms IB, attempts can be made to improve solubility during expression by manipulation of bacterial growth. Interventions tend to be guided by the hypothesis that slowing the rate of expression will allow the protein a longer time to fold. Common approaches include lowering the induction temperature;[351] growing the bacteria in minimal medium; or reducing the concentration of inducing agent. Genetic methods to encourage solubility include the knockout of genes apparently promoting aggregation,

e.g., GroEL,[352] and co-expression with certain molecular chaperones, e.g., DnaK and ClpB.[353] Interventions that account for the specific properties of the target protein can also promote correct folding. For example, expression of soluble lipoic acid synthase (an [Fe-S] cluster containing protein) was enhanced by co-expression of IscS.[354]

5.1.4.2. Solubilisation by modification of the protein sequence.

The most common modification to aid solubility is genetic fusion with a highly soluble protein tag, e.g., N-utilising factor A (NusA); glutathione *S*-transferase (GST); or maltose binding protein (MBP).[355] However, solubilisation may also require direct mutation of the target protein sequence. With the exception of certain critical regions (e.g., the active site), protein function is generally robust to such mutations.[356] However, experimental progress is hampered by the lack of knowledge of the relationship between protein sequence and solubility. Thus, the approach relies on guesses. These can be informed by molecular modelling and comparison with homologous sequences for which expression data are available. Random mutation of the target protein makes fewer *a priori* assumptions, but is not feasible without high throughput technology to screen for soluble mutants. This can be done via genetic fusion to a C-terminal reporter tag, e.g., GFP[357] or CAT,[358] that will only function if the target protein is soluble.

5.1.4.3. Predicting the chance of IB formation.

Given the often-laborious task of solubilising protein by these methods, interest has focussed on predicting the outcome of expression experiments *in silico*, based on primary amino acid sequence. One prominent (but unproven) example is the Wilkinson-Harrison solubility model[359] constructed by examination of the amino acid sequence of

81 over-expressed proteins of known solubility. It was shown that neutrality of charge; and few asparagine, proline, glycine and serine residues (correlating with the propensity to form turns) were the major determinants of solubility. Based on these parameters, a statistical model was generated to predict the probability of achieving soluble protein during over-expression in *E. coli.*

5.1.5. Protein solubilisation *ex vivo.*

5.1.5.1. Solubilisation by denaturation.

Extraction of IB from cells is classically achieved by denaturation with aggressive chaotropes (guanidine hydrochloride and urea).[360] Proteins extracted by these methods must be refolded *in vitro* to a functional state. Levinthal[361] calculated that a protein of 300 amino acids could adopt 10^{300} different folds. Achieving a native fold by refolding appears improbable ($p = 10^{-300}$)! Nevertheless, denatured proteins can be refolded *in vitro*.[362] Examination of the literature shows that refolding of proteins to an active state is often achieved.

Uncovering the most appropriate refolding conditions is an empirical process.[363] It is typically done by exchange of protein into buffer lacking denaturant (by dialysis, dilution or rapid buffer exchange). It is important to note that while it is sometimes possible to refold a protein into a soluble state by these methods, it can be difficult to achieve a functional fold. Hence, the process is most usefully performed using a large-scale screen to maximise the chance of a hit, e.g., with a series of solutions containing different concentrations of buffer, cations and other additives. Temperature, pH and protein concentration are also important parameters.

5.1.5.2. Predicting the outcome of protein refolding experiments.

It would be useful to predict the outcome of refolding experiments *in silico* (under a given set of chemical conditions) based on primary amino acid sequence alone. To my knowledge, only one such algorithm exists, namely, Ho and Middelberg.[364] It was suggested that the tendency to adopt a native fold is inversely related to the rate of protein aggregation, and thus may be considered a function of the second virial coefficient (SVC). Thus, conditions that increase the SVC are likely to increase folding yield. Correlation between measured SVC and the molecular weight and hydropathicity was found; hence, hypothetical SVC may be calculated based on protein sequence. By this method, protein can be classed as less aggregative ($SVC > 2 \times 10^3$ mL.mol.g^{-2}) or more aggregative ($SVC < 2 \times 10^3$ mL.mol.g^{-2}).

5.1.6. Chapter objective and overview.

The objective of the work reported in this chapter was to produce soluble and enzymatically active PfSufS. Following the Materials and Methods, the results section begins with a brief discussion of the properties of the PfSufS sequence as they relate to suitability for heterologous expression. A series of PfSufS expression constructs and the experiments conducted with them are described.

5.2. MATERIALS AND METHODS.

5.2.1. Computational techniques.

Molecular structures were visualised using iMol while protein sequence properties were analysed using SeqVu. Sequences were aligned using the ClustalW algorithm[249]* and refined by eye. The ExPASy ProtParam tool† was used to predict M_r, isoelectric point (pI), molar extinction coefficient at 280 nm (ε),[333] grand average of hydropathy (GRAVY)[233] and instability index (II).[396] Wilkinson-Harrison probabilities of solubility[359] were calculated using an Internet-based algorithm.‡ Stability of proteins in *E. coli* cytosol was predicted by reference to the N-end rule.[397] Second virial coefficients (SVC) were calculated using a Microsoft Excel spreadsheet according to the method of Ho and Middelberg.[364] Transmembrane regions were predicted with the DAS-TMfilter server.[398]¶ N-linked glycosylation sites were predicted using the NetNGlyc 1.0 server.§ O-linked glycosylation sites were predicted using the DictyOGlyc 1.1 server.[399]|| Quality of PCR primers was checked with the Amplify software. DNA sequences were assembled using the ABI Prism Sequencing Analysis software (Perkin Elmer) running on an Apple Power Mac G4.

5.2.2. Nucleic acid techniques.

Nucleic acid techniques were as described (chapter 2). Primer sequences are indicated in Table 5.1.

---

* http://www.ebi.ac.uk/clustalw/
† http://www.expasy.org/tools/protparam.html
‡ http://www.biotech.ou.edu/
¶ http://mendel.imp.univie.ac.at/sat/DAS/DAS.html
§ http://www.cbs.dtu.dk/services/NetNGlyc/
|| http://www.cbs.dtu.dk/services/DictyOGlyc/

TABLE 5.1.

PCR PRIMERS USED IN CHAPTER 5.

ID	DNA sequence
2.3	5'GAGCTCGATTATTTTAAAAATGTAAGAGAACAT
2.4	3'CATTAGTTATTATATTACTACCAT
2.5	5'GTGCATCACTAGGTACAGGT
2.6	3'CTCGAGTTTTTCATTTTTCATTTCATTTAACAT
5.1	3'TTGTGGGCCGAGTTGAACAAAATGATTTTTCAT
5.2	5'AGGCTTGGCATATTACCATTATGGTCAAATACA
5.3	3'TGGTAATATGCCAAGCCTGTTTTGTGGGCCGAGTTGAAC
5.4	5'GTTCAACTCGGCCCACAAAACAGGCTTGGCATATTACCA
5.5	5'GGGGGATCCGATTATTTTAAAAATGTAAGAGAACAT
5.6	3'GGGGAATTCTTTTTCATTTTTCATTTCATTTAACAT
5.7	5'GGGGAATTCGATTATTTTAAAAATGTAAGAGAACAT
5.8	3'GGGCTGCAGTTATTTTTCATTTTTCATTTCATTTAACAT
5.9	5'GGGGAATTCCTGGAAGTTCTGTTCCAGGGGCCCGATTATTTTAAAAATGTAAGAGAACAT
5.10	3'GGGCTGCAGTTAGTGGTGGTGGTGGTGGTGTTTTTCATTTTTCATTTCATTTAACAT
5.11	3'GGGCTGCAGTTAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGTTTTTCATTTTTCATTTCATTTAACAT

5.2.3. Protein techniques.

5.2.3.1. Properties of expression vectors.

Expression vector pTrcHis2-TOPO was obtained from Invitrogen; pET28a(+) and pET43.1b(+) from Novagen; pGEX-6P-1 from Amersham; and pMALc2 and pMALp2 from NEB. pET28a(+) encodes a cassette conferring resistance to kanamycin; the other vectors encode a ß-lactamase conferring resistance to ampicillin. Transcription from pTrcHis2-TOPO is programmed by the *Trc* promoter; whereas transcription from pET28a(+) and pET43.1b(+) is under the control of the T7 promoter. Expression from pGEX-6P-1 is programmed by *Ptac*. In all cases, expression is regulated by the *lac* system and is therefore induced with isopropylthio-ß-D-galactoside (IPTG). Additionally, two expression vectors for use in *Saccharomyces cerevisiae* (pMG and pMH903; Fig. 5.23A) were obtained as gifts from Dr. S. G. Sedgwick (NIMR). Expression from these vectors is programmed by *Pgal* (responsive to galactose). Fusion proteins expressed from all the vectors are depicted in Fig. 5.2, 5.7 and 5.11.

5.2.3.2. Expression culture.

Expression plasmids were transformed into Rosetta-gami (DE3) pLysS (Novagen) or BL21-CodonPlus (DE3)-RIL (Stratagene) competent *E. coli* cells. Rosetta-gami strains are deficient in glutathione reductase and thioredoxin B, thus promoting disulphide bond formation in the cytosol. pLysS and RIL plasmids encode rare codons thereby correcting for the AT bias of *P. falciparum* DNA. Bacteria were grown in Luria-Bertani broth (10 g bactotryptone, 5 g yeast extract, 10 g NaCl to 1 L in redistilled water) or Terrific broth (10 g tryptone, 5 g yeast extract, 10 g NaCl, 6.3 g K_2HPO_4, 1.8 g KH_2PO_4,

sodium citrate 0.45 g, $MgSO_4.7H_2O$ 0.09 g, $(NH_4)_2SO_4$ 0.9 g, glycerol 44 g to 1 L in redistilled water). In addition to antibiotic (ampicillin or kanamycin) required to maintain the expression plasmid (section 5.2.3.1), bacteria harbouring pLysS and RIL plasmids were grown with chloramphenicol. Expression media were routinely enriched with 50 μM pyridoxal 5′-phosphate (PLP). Protein expression experiments in *S. cerevisiae* were conducted by Dr. S. G. Sedgwick.

5.2.3.3. Detection and quantification of expressed proteins.

Proteins were detected by denaturing polyacrylamide gel electrophoresis using the NuPAGE system (Invitrogen) followed by Coomassie brilliant blue staining or Western transfer. On Western blots, His fusions were detected by incubation with 1:1000 His-tag monoclonal antibody (Novagen) followed by 1:6000 anti-mouse IgG-HRP (Sigma). S-tag fusions were detected by incubation with 1:5000 S-protein HRP conjugate reagent (Novagen). GST fusions were detected by incubation with 1 $\mu g.mL^{-1}$ mouse anti-GST (Zymed) followed by 1:2000 HRP-rabbit anti-mouse IgG2b (Zymed). MBP fusions were detected by incubation with 1:10000 rabbit anti-MBP antisera (NEB) followed by 1:3000 goat anti-rabbit IgG (H+L)-HRP conjugate (Bio-Rad). Cross-reactive bands were revealed by ECL (Amersham). Images of gels were captured using an Epson Perfection 2480 Photo scanner and molecular weights of proteins were estimated by comparison to BenchMark, Mark12 or SeeBlue standards (Invitrogen) using image analysis (UVP LabWorks). N-terminal sequencing was done by Babraham Bioscience Technologies (UK). Protein concentrations in solution were measured by absorbance of aromatic amino acids at 280 nm or by BCA protein assay (Pierce) using BSA as the standard. Protein function was assessed by PLP binding and methylene blue (MB^+) assay as described (chapter 4).

5.2.3.4. Liquid chromatography.

Liquid chromatography was conducted on fast protein liquid chromatography (FPLC) apparatus: Äkta Prime or Äkta FPLC (Amersham). Operation of the Äkta FPLC was controlled by the UNICORN software (Amersham) running on a Compaq DESKPRO PC. Buffers intended for use in the high-pressure system of the FPLC were filtered through a polyethersulfone membrane of 0.2 μm pore size.

5.2.4. Engineering of expression constructs.

5.2.4.1. Recovery of mature P. falciparum *sufS.*

P. falciparum sufS (PlasmoDB PF07_0068) predicted to encode the mature enzyme (PfSufS[M]; D62-K546) was recovered from cDNA using the splicing by overlap extension (SOE)-PCR technique of Horton, *et al.*[400] Briefly, PfsufS[M] was split into two overlapping segments. Primers flanked with complementary sequence were designed to amplify each segment such that the two segments could be joined in a second round of PCR. This approach was necessary owing to the failure to recover PfsufS[M] in one PCR reaction (chapter 2).

In the first round of PCR, the 5′ segment was amplified with primers 2.3 and 2.4; while the 3′ segment was amplified with primers 2.5 and 2.6. Thermocycling conditions were: denaturation at 94 °C for 1 min; followed by 35 cycles of 98 °C for 20 s; 45 °C for 30 s; and 72 °C for 1 min. PCR products were gel purified and assembled with a second round of PCR directed by the extreme 5′ and 3′ primers (2.3 and 2.6). Thermocycling conditions were identical to the first round except for a higher annealing temperature

(65 °C). After gel purification, 5′ overhangs were generated by heating the DNA at 72 °C for 10 min with dNTPs (2.5 mM each), 1 × PCR buffer I (Perkin Elmer) and 1 U AmpliTaq (Roche). Final product was extracted in phenol-chloroform and ligated into pCRII (Invitrogen) by TA cloning. Insert was checked by restriction digest and Sanger sequencing. The construct was termed pCRII:PfsufS[M].

5.2.4.2. Construction of chimaeric P. falciparum *sufS.*

A chimaeric version of PfsufS[M] (PfsufS[M]χ) was also generated. In this construct, sequence encoding the PrP-like loop was replaced with a cassette (GGCCCACAAAACAGGCTT; GPQNRLG) modelled to ease folding and derived from the EcSufS sequence (referred to as the *E. coli* bridge). SOE-PCR was employed to generate PfsufS[M]χ in three distinct rounds of PCR.

In the first step, sequence encoding PfSufS[M] was split into three segments that were amplified in three separate PCR reactions from a cDNA template. The segment 5′ distal to the PrP-like loop (N-terminal region) was amplified with primers 2.3 and 2.4. The segment 5′ proximal to the PrP-like loop (middle region) was amplified with primers 2.5 and 5.1. The segment 3′ to the PrP-like loop (C-terminal region) was amplified with primers 5.2 and 2.6. Thermocycling conditions were as described (section 5.2.4.1) with an annealing temperature of 45 °C. PCR products were gel purified.

In order to insert the *E. coli* bridge, products encoding the middle and C-terminal regions were used as targets for a second round of PCR in two distinct reactions. In the first reaction, the middle region was joined to the *E. coli* bridge using primers 2.5 and 5.3. In the second reaction, the C-terminal region was joined to the *E. coli* bridge using

primers 5.4 and 2.6. Touchdown thermocycling conditions for both reactions were: denaturation at 94 °C for 1 min; followed by five cycles of 98 °C for 20 s; 45 °C for 30 s; and 72 °C for 30 s; followed by 30 cycles of 98 °C for 20 s; 60 °C for 30 s; and 72 °C for 1 min.

Products were gel purified before a third round of PCR to join sequence encoding the N-terminal region (made in the first round of PCR) to the *E. coli* bridged middle and C-terminal regions (made in the second round of PCR). This final reaction was directed by the extreme 5′ and 3′ primers (2.3 and 2.6), thus giving PfsufS[M]χ. Thermocycling conditions were as described (section 5.2.4.1), with the exception of the annealing step conducted at 64 °C. Product was screened for the presence of the *E. coli* bridge by sizing on 2 % Metaphor agarose gel (FMC Bioproducts) before insertion into pCRII as described (section 5.2.4.1). Clones were checked for insert by restriction digest and Sanger sequencing. The resulting construct was termed pCRII:PfsufS[M]χ.

5.2.4.3. Cloning into expression vectors.

An expression construct encoding exon 2 of *P. falciparum* SufS (PfsufS[E2]; K98-K546), known as pTrcHis2-TOPO:PfsufS[E2], was obtained from Mr. A. Slatter and Dr. B. Clough. PfsufS[M] and PfsufS[M]χ were recovered by digestion (*Sac*I and *Xho*I) from pCRII:PfsufS[M] and pCRII:PfsufS[M]χ, respectively, and inserted into the corresponding sites in pET43.1b(+), generating pET43.1b(+):PfsufS[M] and pET43.1b(+):PfsufS[M]χ, respectively. PfsufS[M] was additionally cloned into the same restriction sites in pET28a(+), giving pET28a(+):PfsufS[M]. PfsufS[M]χ was recovered from pCRII:PfsufS[M]χ by PCR using primer sets flanked with the appropriate restriction sites for cloning into plasmids pGEX-6P-1, pMALc2, pMALp2,

pMH903 and pMG, generating pGEX-6P-1:PfsufS[M]χ, pMALc2:PfsufS[M]χ, pMALp2:PfsufS[M]χ, pMH903:PfsufS[M]χ and pMG:PfsufS[M]χ, respectively. Primers flanked with appropriate restriction sites (*Bam*HI and *Eco*RI) for insertion into pGEX-6P-1, pMH903 and pMG were 5.5 and 5.6; equivalents flanked with restriction sites (*Eco*RI and *Pst*I) for insertion into pMALc2 and pMALp2 were 5.7 and 5.8. Thermocycling conditions were as described (section 5.2.4.1), with the exception of the annealing step conducted at 57 °C. PCR products were digested with appropriate restriction enzymes and inserted directly into target vectors. Inserts were confirmed by restriction digest and Sanger sequencing.

Additional editing was performed on pMALc2:PfsufS[M]χ to insert a human rhinovirus 3C protease ($3C^{pro}$) recognition site (LEVLFQGP^{401}) between cMBP and PfSufS[M]χ; and His_6 or His_{12} affinity tags at the C-terminus of the fusion protein, generating pMALc2:$3C^{pro}$-PfsufS[M]χ-His_6 and pMALc2:$3C^{pro}$-PfsufS[M]χ-His_{12}, respectively. Briefly, PfsufS[M]χ was recovered by PCR from pMALc2:PfsufS[M]χ using primers flanked with sequence coding for the insertions, i.e., primer 5.9 was combined with either primer 5.10 or 5.11. Thermocycling conditions were as described (section 5.2.4.1), with the exception of the annealing step at 52 °C. PCR products were digested (*Eco*RI and *Pst*I) and ligated into fresh pMALc2 vector. Sequences were confirmed by restriction digest and Sanger sequencing.

5.2.5. Experiments with pTrcHis2-TOPO:PfsufS[E2].

5.2.5.1. Protein expression.

BL21-CodonPlus (DE3)-RIL cells harbouring pTrcHis2-TOPO:PfsufS[E2] were incubated for 16 h in Luria-Bertani broth containing 100 $\mu g.mL^{-1}$ ampicillin. The culture was diluted 1/50 into fresh Luria-Bertani broth containing 50 $\mu g.mL^{-1}$ ampicillin and grown until the optical density at 600 nm (OD600) reached 0.6. Cultures were induced with 1 mM IPTG at 37 °C, 23 °C or 18 °C. Pilot experiments were done in 100 mL of culture medium; while preparative experiments were done in 3 L of culture medium.

5.2.5.2. Protein purification.

For soluble protein preparations, all steps were performed at 4 °C; while for insoluble preparations, steps were performed at room temperature and buffers were supplemented with 6 M urea. Bacterial pellets were recovered by centrifugation and re-suspended in binding buffer (20 mM phosphate, 0.5 M NaCl, 10 mM imidazole, pH 7.6) containing Roche complete EDTA-free protease inhibitor cocktail tablets (1 tablet per 50 mL of suspension). The cell suspension was sonicated on ice with a Sonics and Materials Vibracell probe tip sonicator. Sonicate was centrifuged at 150000*g* for 1 h at 4 °C. The supernatant was applied to an Amersham HiTrap Chelating HP 1 mL column with a Pharmacia LKB Pump P1. Previously the column had been charged with 0.5 mL 0.1 M $NiSO_4$ and equilibrated by flushing with 10 mL binding buffer. Protein was eluted from the column with a linear gradient limited by 1 M imidazole using an Äkta FPLC. Urea

was removed by dialysis for 24 h against buffer containing 50 mM Tris, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT.

5.2.6. Experiments with pET28a(+):PfsufS[M].

Cells harbouring pET28a(+):PfsufS[M] were incubated for 16 h in Luria-Bertani broth containing 100 $\mu g.mL^{-1}$ kanamycin and 34 $\mu g.mL^{-1}$ chloramphenicol. The culture was diluted 1/50 into fresh Terrific broth containing 50 $\mu g.mL^{-1}$ kanamycin, 34 $\mu g.mL^{-1}$ chloramphenicol and 50 μM PLP. At an OD600 of 1.0, the culture was induced for 5 h at 37 °C with 1 mM IPTG. Pilot experiments were done in 100 mL of batch culture.

5.2.7. Experiments with pET43.1b(+):PfsufS[M] and pET43.1b(+):PfsufS[M]χ.

5.2.7.1. Protein expression.

Cells harbouring pET43.1b(+):PfsufS[M] and pET43.1b(+):PfsufS[M]χ were cultured as described (section 5.2.6). Pilot experiments were done in 1 L of batch culture. Conversely, preparative experiments required a culture volume of 40 L and were conducted in a biostat (Large Scale laboratory, NIMR). Cells were harvested by centrifugation and lysed in PBS containing Roche complete EDTA-free protease inhibitor cocktail set (1 tablet per 50 mL) by sonication or passage through a French pressure cell.

5.2.7.2. Protein purification.

Inclusion bodies were prepared essentially as described.[402] Briefly, lysate (section 5.2.7.1) was centrifuged at 12000*g* to recover IB. Pelleted IB was washed repeatedly in PBS and recovered by centrifugation at 12000*g*; prior to a final differential centrifugation at 960*g* to partition the protein component of IB into the supernatant. Supernatant was transferred to another receptacle and IB recovered by centrifugation at 12000*g*. IB was dissolved by gentle stirring for 1 h at 25 °C in 50 mL of buffer containing 20 mM phosphate, pH 7.4, 0.5 M NaCl, 10 mM imidazole, 8 M urea. Ethylenediamine (25 mM) was included in all urea buffers to prevent cyanate accumulation.[403,404] Protein solutions were centrifuged at 150000*g* for 20 min at 4 °C and the supernatant was applied to a 1 mL or 5 mL HisTrap HP column (Amersham) using an Ätka Prime. Protein was eluted by linear gradient limited by 250 mM imidazole.

5.2.7.3. Cleavage experiments.

Fractions containing NusA-PfSufS[M]χ were pooled and concentrated through an Amicon membrane (molecular weight cut-off 10 kDa). Concentrated protein was dialysed for 80 h at 4 °C against 1 L EKMax reaction buffer (50 mM Tris, pH 8.0, 1 mM $CaCl_2$, 0.1 % v/v Tween-20) supplemented with 1 M urea. Protein was centrifuged at 150000*g* to remove precipitate. Soluble protein was digested for 20 h at 4 °C or 37 °C with differing amounts (0.001-4 U) of EnterokinaseMax (Invitrogen).

5.2.7.4. Refolding screen.

NusA-PfSufS[M]χ was submitted to the FoldIt refolding screen (Hampton Research) according to the instructions of the manufacturer. Briefly, fractions containing NusA-PfSufS[M]χ were pooled and dialysed for 16 h at 4 °C against 1 L of buffer containing 100 mM Tris, pH 7.6, 100 mM NaCl, 1 mM DTT, 4 M urea. No precipitation of protein was evident in 4 M urea. Therefore, dialysate was diluted 20-fold into 16 different refolding buffers with addition of PLP in a selection of buffers (Table 5.2). Mixtures were incubated on a rocking platform for 12 h at 4 °C before dialysis in mini dialysis units (molecular weight cut-off 10 kDa) for 4 h at 4 °C (one change of buffer) against 1 L of buffer containing 50 mM NaH_2PO_4, pH 8.0, 150 mM NaCl, in a common container. Dialysate was centrifuged at 150000*g* for 20 min at 4 °C and the supernatants analysed for protein by Western blotting.

5.2.7.5. Preparative refolding.

Concentrated NusA-PfSufS[M]χ (1mL) dissolved in 8 M urea (section 5.2.7.2) was dialysed against 1 L of freshly made refolding buffer (section 5.2.7.4) for 16 h at 4 °C. Refolding buffer was exchanged for 50 mM NaH_2PO_4, pH 8.0, 150 mM NaCl by either a second dialysis for 1 h at 4 °C; or a rapid buffer exchange through a 5 mL HiTrap Desalting column using an Äkta Prime. Soluble protein was recovered by centrifugation at 150000*g* for 20 min at 4 °C.

TABLE 5.2.

COMPOSITION OF BUFFERS IN THE FACTORIAL REFOLDING SCREEN.

NusA-PfSufS[M]χ was submitted to a factorial refolding screen comprising 16 buffers of stated composition. Abbreviations are as follows. PEG 3350: polyethylene glycol 3350; GuHCl: guanidine hydrochloride; GSH: reduced glutathione; GSSG: oxidised glutathione.

FoldIt reagent	Buffer	Salts	PEG 3350	GuHCl	Divalent cation/chelator	Polar/non-polar additives	DTT	Lauryl maltoside	GSH	GSSG	PLP	Solubility
1	50 mM Tris pH 8.2	241 mM NaCl, 10mM KCl	0.05%	0	1.0 mM EDTA	0	0.96 mM	0	0	0	0.48 mM	-
2	50 mM MES pH 6.5	9.65 mM NaCl, 0.40 mM KCl	0	502 mM	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	0	0	0.29 mM	0.96 mM	0.096 mM	0	-
3	50 mM MES pH 6.5	9.65 mM NaCl, 0.40 mM KCl	0.05%	502 mM	1.0 mM EDTA	402 mM Sucrose, 502 mM L-Arginine	0	0	0.96 mM	0.096 mM	0.48 mM	-
4	50 mM Tris pH 8.2	241 mM NaCl, 10mM KCl	0	0	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	402 mM Sucrose, 502 mM L-Arginine	0.96 mM	0.29 mM	0	0	0	+
5	50 mM MES pH 6.5	241 mM NaCl, 10mM KCl	0	0	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	402 mM Sucrose	0	0	0.96 mM	0.096 mM	0.48 mM	-
6	50 mM Tris pH 8.2	9.65 mM NaCl, 0.40 mM KCl	0.05%	502 mM	1.0 mM EDTA	402 mM Sucrose	0.96 mM	0.29 mM	0	0	0	+
7	50 mM Tris pH 8.2	9.65 mM NaCl, 0.40 mM KCl	0	502 mM	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	502 mM L-Arginine	0.96 mM	0	0	0	0.48 mM	-
8	50 mM MES pH 6.5	241 mM NaCl, 10mM KCl	0.05%	0	1.0 mM EDTA	502 mM L-Arginine	0	0.29 mM	0.96 mM	0.096 mM	0	+
9	50 mM MES pH 6.5	241 mM NaCl, 10mM KCl	0.05%	502 mM	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	402 mM Sucrose	0.96 mM	0	0	0	0	-
10	50 mM Tris pH 8.2	9.65 mM NaCl, 0.40 mM KCl	0	0	1.0 mM EDTA	402 mM Sucrose	0	0.29 mM	0.96 mM	0.096 mM	0.48 mM	+
11	50 mM Tris pH 8.2	9.65 mM NaCl, 0.40 mM KCl	0.05%	0	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	502 mM L-Arginine	0	0	0.96 mM	0.096 mM	0	-
12	50 mM MES pH 6.5	241 mM NaCl, 10mM KCl	0	502 mM	1.0 mM EDTA	502 mM L-Arginine	0.96 mM	0.29 mM	0	0	0.48 mM	+
13	50 mM Tris pH 8.2	241 mM NaCl, 10mM KCl	0	502 mM	1.0 mM EDTA	0	0	0	0.96 mM	0.096 mM	0	-
14	50 mM MES pH 6.5	9.65 mM NaCl, 0.40 mM KCl	0.05%	0	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	0	0.96 mM	0.29 mM	0	0	0.48 mM	-
15	50 mM MES pH 6.5	9.65 mM NaCl, 0.40 mM KCl	0	0	1.0 mM EDTA	402 mM Sucrose, 502 mM L-Arginine	0.96 mM	0	0	0	0	-
16	50 mM Tris pH 8.2	241 mM NaCl, 10mM KCl	0.05%	502 mM	2.0 mM $MgCl_2$, 2.0 mM $CaCl_2$	402 mM Sucrose, 502 mM L-Arginine	0	0.29 mM	0.96 mM	0.096 mM	0.48 mM	+

5.2.8. Experiments with pGEX-6P-1:PfsufS[M]χ.

5.2.8.1. Protein expression.

Cells harbouring pGEX-6P-1:PfsufS[M]χ were cultured as described (section 5.2.6). Preparative expressions were done in 3-6 L of culture medium in a batch format.

5.2.8.2. Protein solubilisation with N-*laurylsarcosine.*

Protein was solubilised with *N*-laurylsarcosine essentially as described.[383] Bacterial pellets were washed in cold STE (10 mM Tris, pH 8.0, 150 mM NaCl, 1 mM EDTA) prior to re-suspension in STE containing 100 μg.mL^{-1} chicken egg white lysozyme (Sigma) and Roche complete EDTA-free protease inhibitor cocktail set (2 tablets per 50 mL). The suspension was incubated on ice for 15 min before addition of DTT to a final concentration of 5 mM and *N*-laurylsarcosine to a final concentration of 2 % w/v. For small volumes (500 μL), the suspension was sonicated for 1 min in 10 s bursts immersed in an ice-water mix in an Ultrawave water bath sonicator. For large volumes (25 mL), the suspension was sonicated packed in ice for 1 min in 10 s bursts using a Sonics and Materials Vibracell probe tip sonicator. To recover soluble protein, lysate was centrifuged at 150000*g* for 30 min at 4 °C. Supernatant was neutralised to a final concentration of 2 % v/v Triton X-100 before further processing.

5.2.8.3. Pilot scale purification over glutathione-agarose

Lyophilised glutathione-agarose (Sigma) was swollen in redistilled water (200 mL.g^{-1}), washed repeatedly with PBS and equilibrated in PBS containing 2 % v/v Triton X-100.

Supernatant containing the protein (section 5.2.8.2) was added to the glutathione-agarose gel and the mixture incubated on a rotator for 1 h at 4 °C. For batch purifications, washing and elution was done by repeated re-suspension of gel and centrifugation (500*g* for 1 min). For small scale purifications, gel was packed by gravity into a 10 mL Bio-Rad plastic column prior to gravity flow washing and elution. Washing buffer was PBS containing 2% v/v Triton X-100 while eluting buffer contained 10 mM reduced L-glutathione, 75 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM DTT and either 2% v/v Triton X-100 or 2% w/v N-octyl-ß-D-glucopyranoside.

5.2.8.4. Purification by liquid chromatography.

Preparative purifications were done by FPLC at 4 °C. Supernatant containing protein (section 5.2.8.2) was applied to a 1 mL GSTrap HP column (Amersham) at a flow rate of 0.2 $mL.min^{-1}$ using an Äkta FPLC. The column was washed at a flow rate of 1 $mL.min^{-1}$ in three steps. The first and third washes were done with 10 mL PBS containing 2% v/v Triton X-100; while the second wash was done with 20 mL washing buffer. Protein was eluted at a flow rate of 1 $mL.min^{-1}$ along a linear gradient (0-10 mM L-glutathione reduced) in buffer containing 75 mM HEPES, pH 7.4, 150 mM NaCl, 5 mM DTT and either 2% v/v Triton X-100 or 2% w/v N-octyl-ß-D-glucopyranoside.

5.2.8.5. Cleavage experiments.

Fractions containing GST-PfSufS[M]χ were pooled and buffer exchanged with PreScission cleavage buffer (50 mM Tris, pH 7.0, 150 mM NaCl, 1 mM EDTA, 1 mM DTT) by passage through a HiTrap Desalting column using an Äkta Prime. Protein was treated with 160 U of a GST-tagged version of $3C^{pro}$ (Amersham PreScission Protease)

and digestion allowed to proceed for 24 h at 4 °C; samples were taken at regular intervals and snap frozen in N_2 (l).

5.2.9. Experiments with pMALc2:PfsufS[M]χ and pMALp2:PfsufS[M]χ.

5.2.9.1. Protein expression.

Cells harbouring pMALc2:PfsufS[M]χ and pMALp2:PfsufS[M]χ were cultured as described (section 5.2.6), except that D-glucose (0.2% w/v) was added to suppress amylase expression; and cultures were induced for 16 h at 14 °C with 0.3 mM IPTG. Preparative expressions were done in 3-6 L of culture medium in a batch format.

5.2.9.2. Pilot scale purification.

Bacterial pellets were re-suspended in "BugBuster" protein extraction reagent containing 25 $U.mL^{-1}$ Benzonase nuclease, 1 $KU.mL^{-1}$ chicken egg white lysozyme, 1 mM DTT, 1mM EDTA and Roche EDTA-free protease inhibitor cocktail set (2 tablets per 50 mL). The suspension was incubated at 25°C for 20 min on a rocking platform before centrifugation at 16000*g* for 20 min at 4°C. Supernatant was mixed with pre-swollen amylose resin (NEB) that had been repeatedly washed with 20 mM Tris, pH 7.4, 0.2 M NaCl, 1 mM DTT, 1 mM EDTA. The mixture was incubated on a rotator for 20 h at 4°C before the resin was packed by gravity into a 10 mL Bio-Rad plastic column. The column was washed with 10 mL washing buffer and eluted with washing buffer containing 10 mM maltose. Elution was also attempted by cleavage with Factor Xa. After protein application and washing, resin was mixed with 400 $\mu g.mL^{-1}$ Factor Xa on a rotator for 80 h at 4 °C.

5.2.9.3. Purification by affinity chromatography.

Protein was purified using an Äkta FPLC through an Omni 10 mm column packed by gravity with 1 mL of amylose resin (NEB). All steps of the purification were done at 4 °C. Bacterial pellets were re-suspended in column buffer containing 20 mM Tris, pH 7.4, 0.2 M NaCl. Roche EDTA-free protease inhibitor cocktail set was added (2 tablets per 50 mL). The suspension was spiked with 1 $KU.mL^{-1}$ chicken egg white lysozyme prior to incubation on ice for 20 min; sonication; and centrifugation at 150000*g* for 30 min at 4 °C. Supernatant was diluted 1/5 with fresh column buffer and applied to the amylose at a flow rate of 0.2 $mL.min^{-1}$. The resin was washed at 1 $mL.min^{-1}$ with column buffer. Protein was eluted at a flow rate of 1 $mL.min^{-1}$ in column buffer with a linear gradient limited by 10 mM maltose. Care was taken to minimise degradation of the amylose column due to amylases in the bacterial lysate. Therefore, a maximum of 50 mL of sample was injected each time. Elutions were chased with 0.1 % v/v SDS followed by redistilled water. After five elutions, the column was re-packed with fresh amylose resin.

5.2.10. Experiments with pMALc2:3C^{pro}-PfsufS[M]χ-His$_{12}$.

Soluble 3C^{pro}-PfSufS[M]χ-His_{12} protein was expressed and extracted as described (sections 5.2.6 and 5.2.9.3) but with some modifications. Critically, protease inhibitors were not used. Exploiting the C-terminal His-tag, soluble protein was purified through a 1 mL HiTrap chelating column (Amersham) charged with Ni^{2+}. Supernatant containing protein (section 5.2.9.2) was applied to the column at a flow rate of 1 $mL.min^{-1}$ using an Äkta Prime. The column was washed at a flow rate of 1 $mL.min^{-1}$ with binding buffer. PreScission Protease (160 U) in a carrier volume of 5 mL binding buffer was injected

into the column at a rate of 0.2 mL.min^{-1}. To allow digestion, the column was sealed and incubated for 20 h at 5 °C. The column was washed extensively in binding buffer and protein was eluted at a flow rate of 1 mL.min^{-1} along a linear gradient limited by 1 M imidazole.

5.2.11. Experiments with pMH903:PfsufS[M]χ.

Vector pMH903:PfsufS[M]χ was transformed into *S. cerevisiae* and protein was expressed using proprietary methods developed Dr. S. G. Sedgwick.

5.3. RESULTS AND DISCUSSION.

5.3.1. Sequence analysis of *P. falciparum* SufS.

5.3.1.1. Codon bias.

A major determinant of low protein expression level in heterologous hosts is the high AT content of *Plasmodium* spp. DNA (review in Withers-Martinez, *et al.*[345]). At the start of this project, only PfsufS (77 % AT) and PysufS (76 % AT) sequences were available in the database. I used the *P. falciparum* gene for my expression experiments, relying on the availability of helper plasmids to supply rare codons. However, other sufS constructs already expressed in *E. coli* in a soluble form have a more equitable AT bias, e.g., EcsufS (GenBank AB055108) (45 % AT) and AtsufS (GenBank AY078068) (55 % AT). Apicomplexan sufS sequences recently deposited in the database include PvsufS (49 % AT); and TgsufS (48 % AT). These might constitute useful future starting points for over-expression.

5.3.1.2. Protein stability.

Protein stability is essential for accumulation of expressed protein. The N-end rule indicates that proteins are stable in the *E. coli* cytosol (Table 5.3). Stability *in vitro* is also important, allowing the protein to be manipulated at the bench. By reference to the instability index (II), PfSufS was predicted to be unstable (II > 40). However, addition of fusion partners greatly improved the predicted *in vitro* stability (Table 5.3).

TABLE 5.3.

BIOINFORMATIC SUMMARY OF SUFS PROTEINS.

The properties of SufS proteins described in this thesis were determined *in silico*. Briefly, the ExPASy ProtParam tool was used to predict M_r, isoelectric point (pI), molar extinction coefficient at 280 nm (ε), grand average of hydropathy (GRAVY) and instability index (II). The Wilkinson-Harrison score for each protein was determined and reported as a chance of obtaining soluble protein during over-expression in *E. coli*. Stability of proteins in *E. coli* cytosol ($t_{1/2}$) was predicted by reference to the N-terminal amino acid (N-end). Second virial coefficients (SVC) were calculated using a Microsoft Excel spreadsheet.

Protein	M_r (kDa)	pI	ε ($M^{-1}.cm^{-1}$)	GRAVY	N-end	$t_{1/2}$	II	Chance of solubility (%)	SVC ($mL.mol.g^{-2}$) × 10^3
PfSufS	65	9.44	67590	-0.44	M	>10 h	44.15	24.2	0.36
PfSufS[E2]	53	9.4	59670	-0.482	K	2 min	44.34	21.3	0.94
PfSufS[E2]-cMyc	57	9.24	60950	-0.505	M	>10 h	42.79	16.5	0.95
PfSufS[M]	57	9.35	62350	-0.486	D	>10 h	44.89	20.2	0.83
NusA-PfSufS[M]	119	6.05	92910	-0.437	M	>10 h	39.47	50.9	-1.27
His_6-PfSufS[M]-His_6	62	9.32	62350	-0.54	M	>10 h	43.17	15.4	0.96
PfSufS[M]χ	49	9.31	55950	-0.328	D	>10 h	46.18	18.5	0.24
NusA-PfSufS[M]χ	111	5.77	86510	-0.366	M	>10 h	39.67	55.5	-1.41
GST-PfSufS[M]χ	78	8.91	96870	-0.356	M	>10 h	42.34	17.9	-0.46
cMBP-PfSufS[M]χ	92	8.65	120670	-0.364	M	>10 h	34.12	17.3	-0.86
cMBP-$3C^{pro}$-PfSufS[M]χ-His_6	94	8.59	120670	-0.374	M	>10 h	33.47	18.4	-0.86
cMBP-$3C^{pro}$-PfSufS[M]χ-His_{12}	95	8.59	120670	-0.395	M	>10 h	33.3	18.8	-0.77
pMBP-PfSufS[M]χ	95	8.8	120670	-0.324	M	>10 h	33.19	16.4	-1.14
His_6-EcSufS	48	5.97	49170	-0.16	M	>10 h	33.86	45	-0.60
AtSufS-His_6	48	6.09	37790	-0.052	M	>10 h	30.34	46.4	-1.16

5.3.1.3. Protein solubility.

Attempts to over-express *P. falciparum* proteins in *E. coli* in an active form may also fail due to aggregation. PfSufS protein sequence lacks obvious trans-membrane regions (chapter 2), suggesting that soluble PfSufS might be produced. However, the Wilkinson-Harrison score was low (24.2 %), suggesting that the protein would be insoluble. This contrasts with the higher Wilkinson-Harrison scores of EcSufS (45 %) and AtSufS (46.4 %).

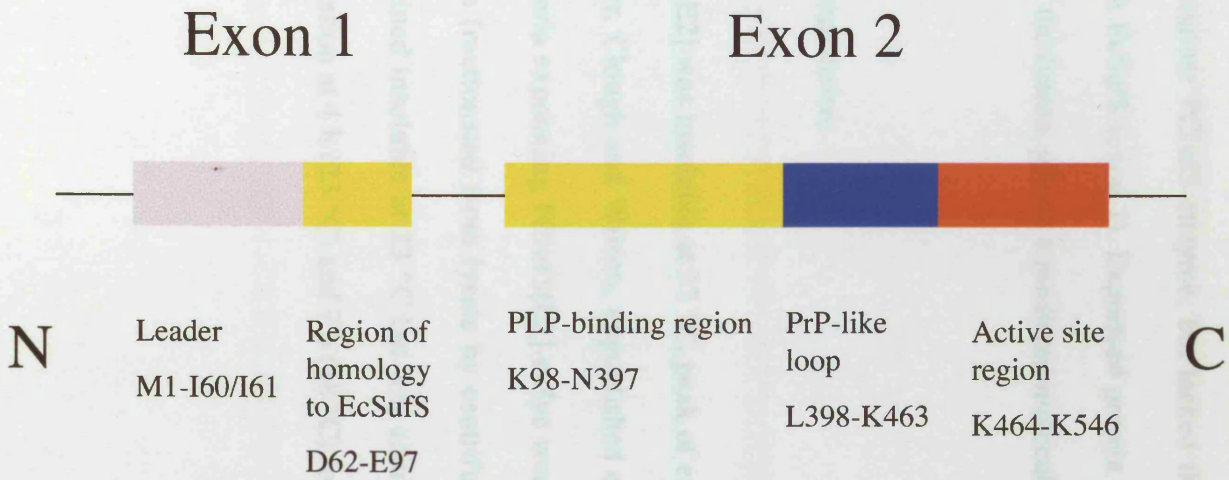
Given such potential for insolubility, PfSufS protein sequence was modified to promote solubility. For example, expression of soluble AtSufS was achieved without the chloroplast-targeting signal. Hence, the bipartite plastid-targeting signal of PfSufS (M1-I61; chapter 3) was absent in all my expression constructs. In some constructs, further modifications were introduced (section 5.3.5) and solubilising fusion partners were used. Major features of PfSufS discussed in this chapter are indicated schematically in Fig. 5.1.

However, it is important to realise that none of these interventions produced a significant impact on the Wilkinson-Harrison score, with the exception of the NusA fusion (Table 5.3). Finally, it is worth noting the aggregative SVC values were exhibited by all the constructs, suggesting a tendency to precipitate *in vitro*. (Table 5.3).

FIGURE 5.1.

SCHEMATIC REPRESENTATION OF *P. FALCIPARUM* SUFS SHOWING MAJOR FEATURES DISCUSSED IN CHAPTER 5.

Amino acids corresponding to indicated regions are reported in the one letter code.


Exon 1
Exon 2
N
Leader
M1-I60/I61
Region of homology to EcSufS
D62-E97
PLP-binding region
K98-N397
PrP-like loop
L398-K463
Active site region
K464-K546
C

5.3.2. Experiments with pTrcHis2-TOPO:PfsufS[E2].

5.3.2.1. Properties of the expression construct.

Vector pTrcHis2-TOPO:PfsufS[E2], encoding PfSufS[E2]-cMyc included PLP-binding and active sites of the mature PfSufS enzyme, but lacked the leader sequence and a region of homology with EcSufS (exon 1). Expressed protein contained cMyc and His_6 tags at the C-terminal of the fusion, giving a predicted molecular weight of 57 kDa (Fig. 5.2).

5.3.2.2. Protein expression.

Over-expressed PfSufS[E2] was insoluble at 37 °C; peak of expression was achieved 3 h after induction (Slatter, Clough and Wilson, unpublished observations). In order to promote solubility, bacteria expressing PfSufS[E2]-cMyc were grown at 23 °C and 18 °C. Soluble protein was fractionated from lysate by centrifuging at 150000*g*. It was found that protein remained insoluble at 23 °C but was soluble at 18 °C (Fig. 5.3). I noted a peak of accumulation at 4 h (23 °C) and 7 h (18 °C) followed by sudden (1-2 h) degradation (Fig. 5.3).

FIGURE 5.2.

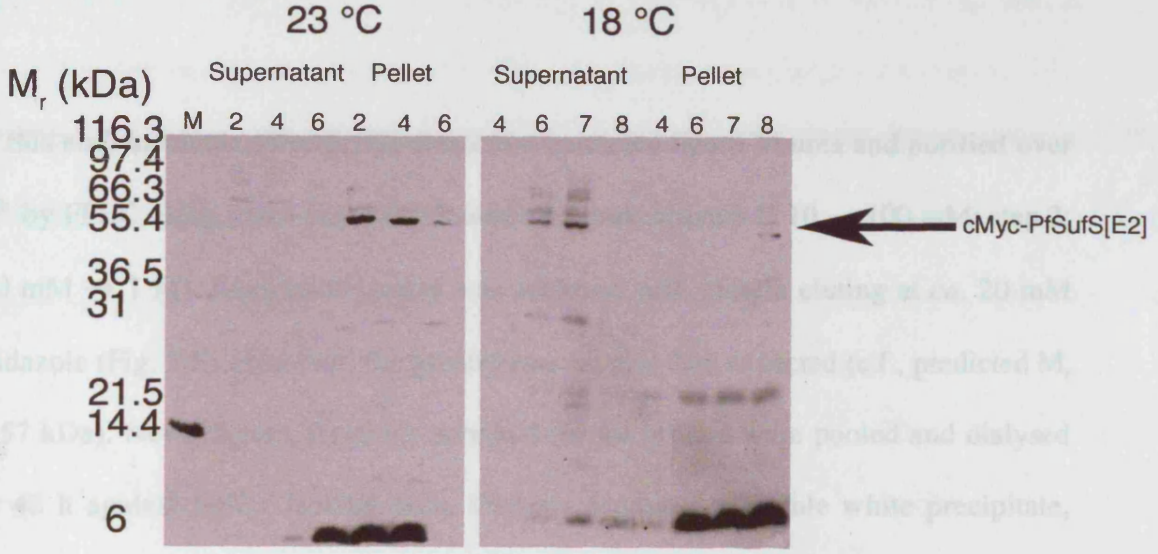
SCHEMATIC OF PFSUFS[E2] FUSION PROTEIN.


N
C
PfSufS[E2]-cMyc
57 kDa
PfSufS[E2]
cMyc
His6

FIGURE 5.3.

EXPRESSION OF PFSUFS[E2]-cMYC AT LOW TEMPERATURE PROBED BY WESTERN BLOT.

E. coli cells harbouring pTrcHis2-TOPO:PfsufS[E2] were induced with 1 mM IPTG at 23 °C and 18 °C. Samples were taken at intervals following induction. Protein was extracted with "BugBuster" reagent (Novagen) and the lysate centrifuged (150000*g*) to separate soluble (supernatant) and insoluble (pellet) protein. Protein was resolved through NuPage 12 % BT gels and detected with anti-His monoclonal antibody by Western blot. Evidence for soluble expression of the fusion protein after 6-7 h induction at 18 °C is indicated with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); 2, 4, 6, 7 and 8: 2, 4, 6, 7 and 8 h post-induction.


23 °C
18 °C
Supernatant
Pellet
Supernatant
Pellet
M r (kDa)
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
M 2 4 6 2 4 6
4 6 7 8 4 6 7 8
cMyc-PfSufS[E2]

5.3.2.3. Affinity purification.

Attempts were made to purify soluble PfSufS[E2]-cMyc produced from 3 L of culture induced for 7 h at 18 °C. Affinity purification over Ni^{2+} was done with a linear imidazole gradient using FPLC (Fig. 5.4). However, purification did not significantly enrich for PfSufS[E2]-cMyc, probably due to the low level of expression of soluble protein. It is important to realise that protein could not be enriched during culture due to instability after 7 h of induction (section 5.3.2.2). Hence I decided to abandon low temperature methods and instead focus on insoluble PfSufS[E2]-cMyc expressed at 37 °C.

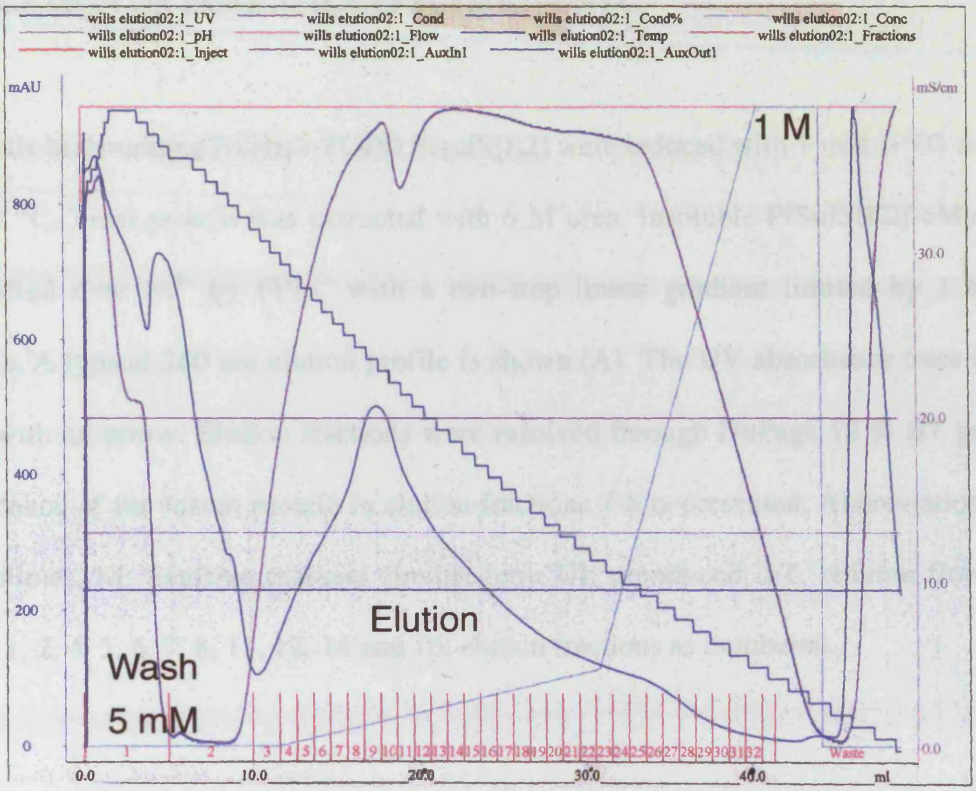
To this end, insoluble PfSufS[E2]-cMyc was extracted into 6 M urea and purified over Ni^{2+} by FPLC using a two-step linear imidazole gradient (step 1: 10 → 100 mM; step 2: 100 mM → 1 M). Reasonable purity was achieved with protein eluting at *ca.* 20 mM imidazole (Fig. 5.5). However, the protein was smaller than expected (c.f., predicted M_r of 57 kDa). Nevertheless, fractions enriched for the protein were pooled and dialysed for 48 h against buffer lacking urea. Dialysis produced a visible white precipitate, possibly representing PfSufS[E2]-cMyc. Further diminution in the size of the protein was noted, suggesting truncation or degradation. Therefore, I performed Western blotting using commercially available anti-cMyc; and anti-cMyc monoclonal antibody generated in-house by Dr. M. J. Blackman (NIMR). This technique indicated significant truncation of the fusion protein (Fig. 5.6).

FIGURE 5.4.

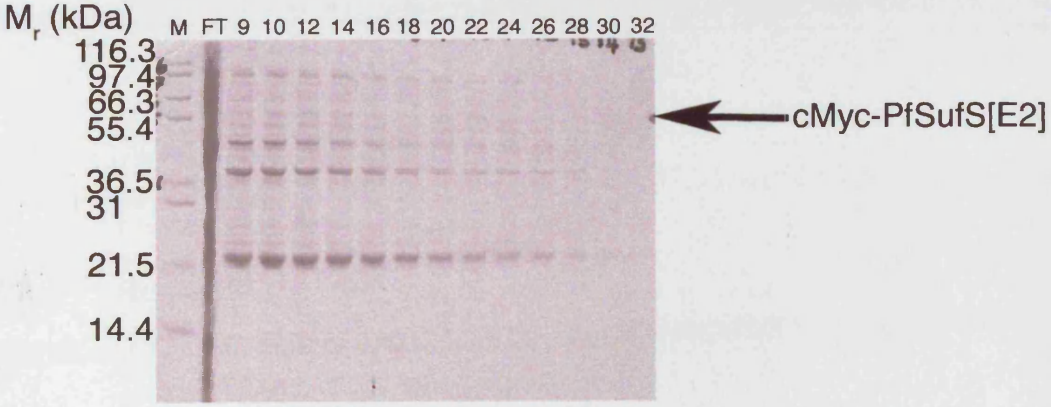
PURIFICATION OF SOLUBLE PFSUFS[E2]-cMYC.

E. coli cells harbouring pTrcHis2-TOPO:PfsufS[E2] were induced with 1 mM IPTG for 7 h at 18 °C. Soluble protein was extracted with "BugBuster" reagent (Novagen). Soluble PfSufS[E2]-cMyc was purified over Ni^{2+} by FPLC with a two-step linear gradient limited by 1 M imidazole. A typical 280 nm elution profile is shown (A). The UV absorbance trace is marked with an arrow. Elution fractions were resolved through NuPage 12 % BT gel (B) and detected with His-tag monoclonal antibody (Novagen) on Western blot (C). Evidence of the fusion protein in elution fractions 30-32 is presented (C). Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); FT: column flow-through; 9, 10, 12, 14, 16, 18. 20, 22, 24, 26, 28, 30 and 32: elution fractions as numbered.

A


wills elution02:1_UV
wills elution02:1_Cond
wills elution02:1_Cond%
wills elution02:1_Conc
wills elution02:1_pH
wills elution02:1_Flow
wills elution02:1_Temp
wills elution02:1_Fractions
wills elution02:1_Inject
wills elution02:1_AuxIn1
wills elution02:1_AuxOut1
mAU
mS/cm
1 M
Elution
Wash
5 mM
Waste
ml

B


M_r (kDa)
M FT 9 10 12 14 16 18 20 22 24 26 28 30 32
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMyc-PfSufS[E2]

C

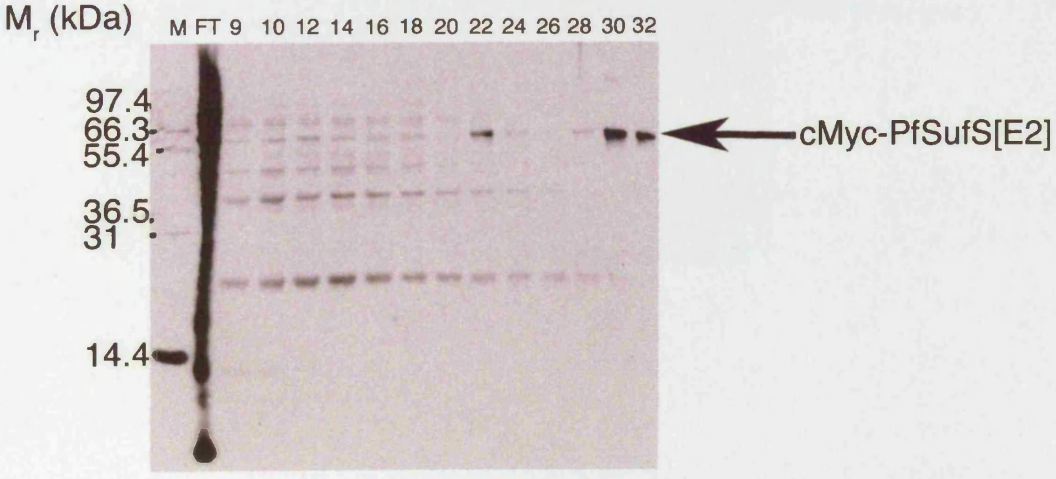
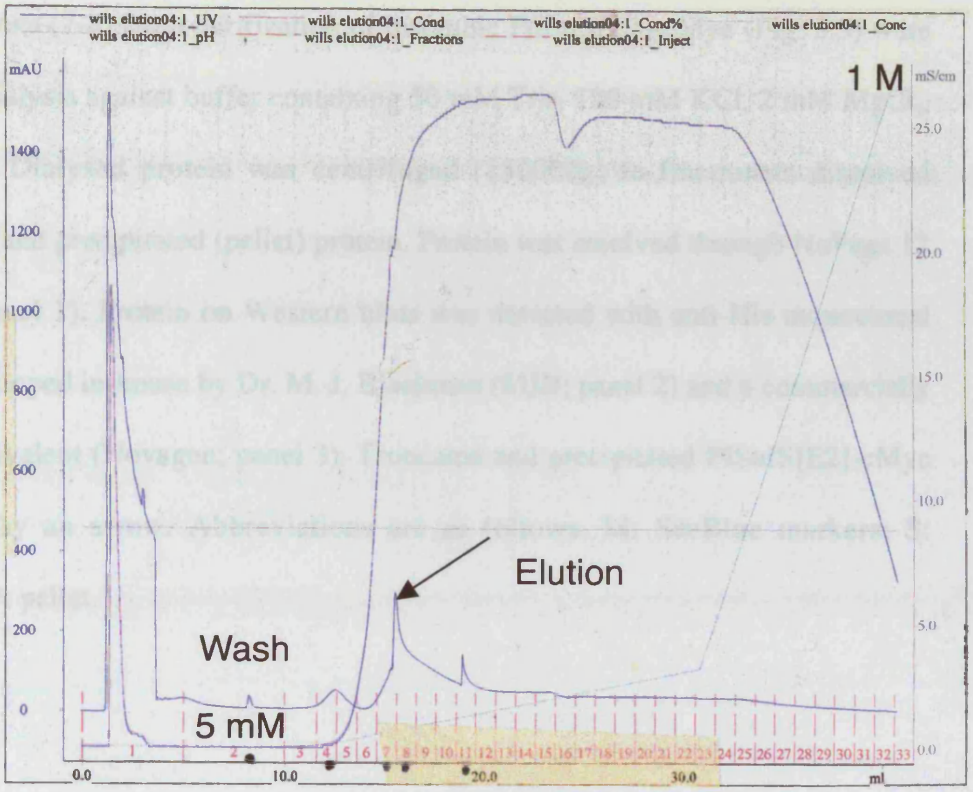

M_r (kDa)
M FT 9 10 12 14 16 18 20 22 24 26 28 30 32
97.4
66.3
55.4
36.5
31
14.4
cMyc-PfSufS[E2]

FIGURE 5.5.

PURIFICATION OF INSOLUBLE PFSUFS[E2]-cMYC.

E. coli cells harbouring pTrcHis2-TOPO:PfsufS[E2] were induced with 1 mM IPTG for 3 h at 37 °C. Total protein was extracted with 6 M urea. Insoluble PfSufS[E2]-cMyc was purified over Ni^{2+} by FPLC with a two-step linear gradient limited by 1 M imidazole. A typical 280 nm elution profile is shown (A). The UV absorbance trace is marked with an arrow. Elution fractions were resolved through NuPage 12 % BT gel (B). Evidence of the fusion protein in elution fractions 7-8 is presented. Abbreviations are as follows. M: SeeBlue markers (Invitrogen); UI: uninduced; FT: column flow-through; 1, 2, 4, 5, 6, 7, 8, 11, 12, 14 and 16: elution fractions as numbered.

A


wills elution04:1_UV
wills elution04:1_pH
wills elution04:1_Cond
wills elution04:1_Fractions
wills elution04:1_Cond%
wills elution04:1_Inject
wills elution04:1_Conc
mAU
mS/cm
1 M
Elution
Wash
5 mM
ml

B

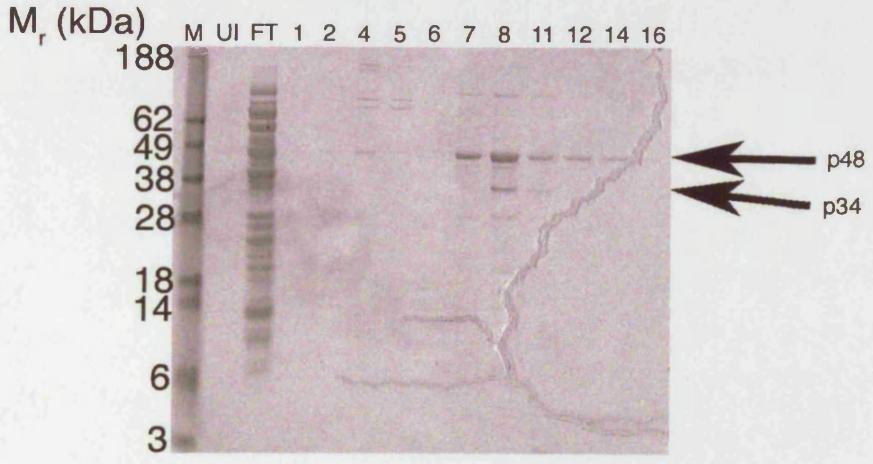
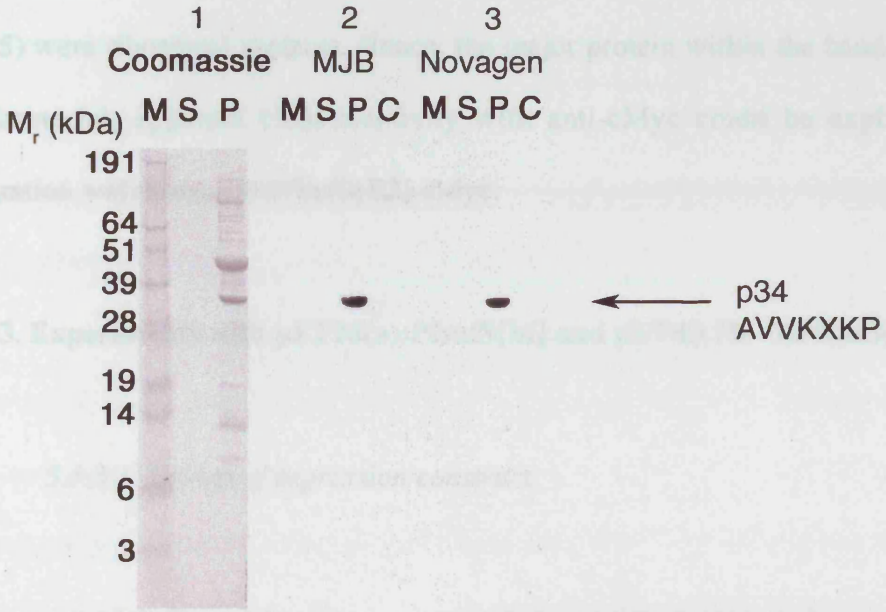

M_r (kDa)
M UI FT 1 2 4 5 6 7 8 11 12 14 16
188
62
49
38
28
18
14
6
3
p48
p34

FIGURE 5.6.

IMMUNO-REACTIVITY OF PFSUFS[E2]-cMYC.

Elution fractions (7-8) from purification of insoluble PfSufS[E2]-cMyc (Fig. 5.5) were refolded by dialysis against buffer containing 50 mM Tris, 100 mM KCl, 2 mM $MgCl_2$, 1 mM DTT. Dialysed protein was centrifuged (150000*g*) to fractionate dissolved (supernatant) and precipitated (pellet) protein. Protein was resolved through NuPage 12 % BT gel (panel 1). Protein on Western blots was detected with anti-His monoclonal antibody developed in-house by Dr. M. J. Blackman (MJB; panel 2) and a commercially available equivalent (Novagen; panel 3). Truncated and precipitated PfSufS[E2]-cMyc is indicated by an arrow. Abbreviations are as follows. M: SeeBlue markers; S: supernatant; P: pellet.


1
2
3
Coomassie
MJB
Novagen
M S P
M S P C
M S P C
M$_r$ (kDa)
191
64
51
39
28
19
14
6
3
p34
AVVKXKP

To further investigate the cause of this phenomenon, pTrcHis2TOPO:PfsufS[E2] plasmids were isolated directly from expression culture to be digested and re-sequenced. These data confirmed the integrity of the insert. However, given that the cMyc epitope tag was at the C-terminal of the fusion protein, the possibility existed of degradation; or transcriptional starts at internal sites, leading to expression of protein lacking a complete N-terminal. To test this hypothesis, N-terminal sequencing was conducted. However, the N-terminal motif (AVVKXKP) did not correspond to any part of PfSufS[E2]-cMyc. Therefore, sequence was searched against the NCBI database using the BLASTP algorithm ("short, nearly exact matches"). Closest in similarity (E = 6375) were ribosomal proteins. Hence, the major protein within the band was likely to be bacterial; apparent cross-reactivity with anti-cMyc could be explained by co-migration with truncated PfSufS[E2]-cMyc.

5.3.3. Experiments with pET28(a):PfsufS[M] and pET43.1b(+):PfsufS[M].

5.3.3.1. Design of expression construct.

Alignment of SufS protein sequences revealed that PfSufS[M] likely encompassed *ca.* 36 amino acids (D62-E97) encoded by exon 1 (additional to amino acids encoded by exon 2). Hence, original PfSufS[E2] constructs (section 5.3.2) may have lacked the potential to be enzymatically active. Therefore, the strategy of cloning only exon 2 was modified to include the homologous region of exon 1.

5.3.3.2. Engineering of expression constructs.

PfsufS[M] was successfully constructed using the SOE-PCR technique and ligated into pET28a(+) and pET43.1b(+) vectors.

5.3.3.3. Properties of the fusion proteins.

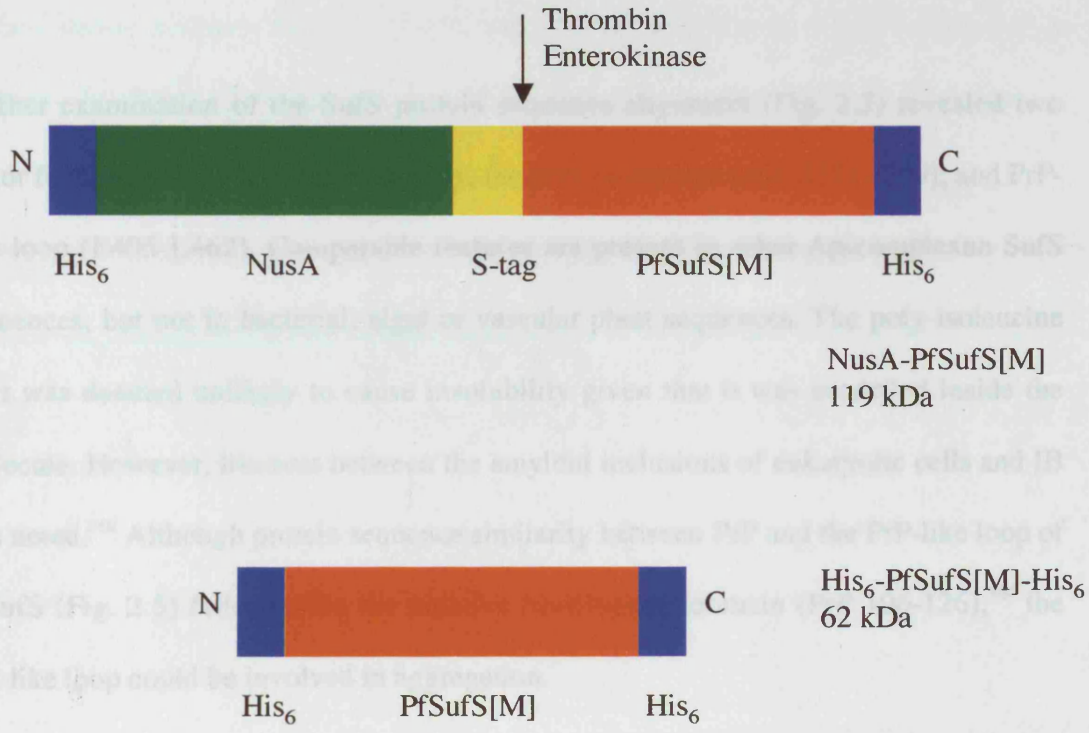
Protein expressed from the pET28a(+) construct lacks solubility tags, possessing only N and C-terminal His_6 tags. Conversely, protein expressed from pET43.1b(+) is designed to contain a large N-terminal extension composed of the highly soluble N-utilising factor A (NusA: GenBank AAC76203); a His_6-tag; an S-tag; and recognition sites for thrombin I and enterokinase; a second His_6-tag is C-terminal to the insert. Schematic representations of fusion proteins are provided (Fig. 5.7).

It is useful to consider the solubilising properties of NusA, an *E. coli* transcriptional anti-terminator protein.[365] Davis, *et al.*[366] suggested that NusA was the most likely *E. coli* protein to promote solubility, as identified from genome-wide screen of *E. coli* proteins using a modified version of the Wilkinson-Harrison solubility model. As expected, NusA markedly increased the Wilkinson-Harrison score (20.2 → 50.9 %) of the PfSufS[M] construct. The tag has since been used to solubilise numerous bacterial, animal and plant proteins over-expressed in *E. coli*.[367] The S-tag additionally permits detection of the fusion protein by interaction with HRP-conjugated S-protein.[368]

FIGURE 5.7.

SCHEMATICS OF PFSUFS[M] FUSION PROTEINS.

PfSufS[M] fusions expressed from pET43.1b(+):PfsufS[M] (A) and pET28a(+):PfsufS[M] (B). Thrombin and enterokinase sites are indicated. Abbreviations are as follows. NusA: N-utilising substance A.


Thrombin
Enterokinase
N
C
His6
NusA
S-tag
PfSufS[M]
His6
NusA-PfSufS[M]
119 kDa
N
C
His6
PfSufS[M]
His6
His6-PfSufS[M]-His6
62 kDa

5.3.3.4. Protein expression.

Experiments done with these constructs revealed a lack of convincing evidence for induction from pET28a(+):PfSufS[M]; while protein produced from pET43.1b(+):PfSufS[M] was truncated (Fig. 5.8). Therefore, I conducted no further experiments with these constructs.

5.3.4. Design and construction of PfsufS[M]χ.

Further examination of the SufS protein sequence alignment (Fig. 2.3) revealed two major features unique to PfSufS, namely, the poly-isoleucine tract (I255-I259); and PrP-like loop (E405-L462). Comparable features are present in other Apicomplexan SufS sequences; but not in bacterial, algal or vascular plant sequences. The poly-isoleucine tract was deemed unlikely to cause insolubility given that it was modelled inside the molecule. However, likeness between the amyloid inclusions of eukaryotic cells and IB was noted.[348] Although protein sequence similarity between PrP and the PrP-like loop of PfSufS (Fig. 2.5) falls outside the putative fibrillogenic domain (PrP 106-126),[369] the PrP-like loop could be involved in aggregation.

To circumvent potential insolubility due to the PrP-like loop, a deletion strategy was adopted. A structure-based alignment of several representative SufS sequences was done by Dr. J. W. Saldanha (Fig. 5.9). The loop was revealed to be longer (66 amino acids; L398-K463) than originally indicated (Fig. 2.3), constituting 12 % of the entire PfSufS ORF.

FIGURE 5.8.

EXPRESSION OF PFSUFS[M].

E. coli cells harbouring pET28a(+):PfsufS[M] and pET43.1b(+):PfsufS[M] were induced with 1 mM IPTG at 37 °C for 3 h. Expressed protein was resolved through NuPage 4-12 % BT gel (panel 1) and detected with anti-His antibody on Western blot (panel 2). Expression of NusA-PfSufS[M] was additionally probed with S-protein HRP (panel 3). NusA-PfSufS[M] was expressed in a truncated form (*ca.* 92 kDa). There was no convincing evidence for induction of protein from pET28a(+). Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); UI: uninduced control; I: induced; 28: pET28a(+):PfsufS[M]; 43: pET43.1b(+):PfsufS[M].

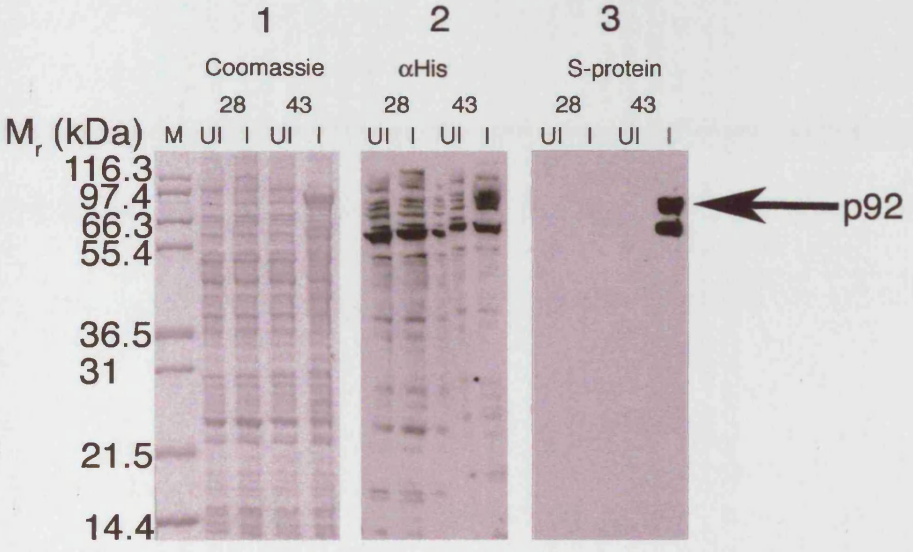

1
2
3
Coomassie
αHis
S-protein
28
43
28
43
28
43
M
UI
I
UI
I
UI
I
UI
I
UI
I
UI
I
M_r (kDa)
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
p92

FIGURE 5.9.

STRUCTURE-BASED ALIGNMENT OF SUFS PROTEIN SEQUENCES.

In order to delineate the PrP-like loop of PfSufS, a rigorous structure-based alignment of SufS protein sequences was done by Dr. J. W. Saldanha. Alignment between the conserved lysine binding (SXHK) and active site (GXHCX) motifs is presented. The PrP-like loop of PfSufS is underlined. Abbreviations are as follows. ECOLI: EcSufS; STYPH: *Salmonella typhimurium* SufS; YPEST: *Yersinia pestis* SufS; ECHRY: *E. chrysanthemi* SufS; ATHAL: AtSufS; MTUBE: *M. tuberculosis* SufS; PFALC: PfSufS; PKNOW: PkSufS; PYOEL: PySufS; TGOND: TgSufS.

```
ECOLI    SGHKLYGPTGIGILYVKEALLQEMP--PWEGGGSMIATVSLSEGTTWTKAPWRFEAGTPNTGGIIGLGAALEYVSALG----LNNIAEYEQNLMHYALSQLES-VPDLTLYGPQ------
STYPH    SGHKMLGPNGIGVLYGRRELLAQMP--PFLTGGSMIETVTM-EGATYAPAPQRFEAGTPMTSQVVGLAAAARYLGAIG----MAAVEAHERELVAAAIEGLSG-IDGVRILGPTS-----
YPEST    SGHKLYGPSGIGILYGKSALLQQMP--PWEGGGAMIKTVSLTQGTTFADAPWRFEAGSPNTAGIMGLGAAIDYVTELG----LLPIQQYEQSLMHYALAQLSQ-IKSLTLYGPT------
ECHRY    SGHKLYGPTGIGILYVKEALLQEMP--PWEGGGSMISTVSLTQGTTWAKAPWRFEAGTPNTGGIIGLGAAIDYVTSLG----LDKIGDYEQMLMRYALEQLAQ-VPDITLYGPA------
MTUBE    SGHKLYGPSGIGVLYGKSELLQAMP--PWEGGGAMIREVSLTQGTTYADPPWRFEAGSPHVAGIIGLGAALDYVSALG----VDAIQAHEGLLMRYALASLAE-VPTLRLYGPV------
ATHAL    SSHKMCGPTGIGFLYGKSDLLHSMP--PFLGGGEMISDVFL-DHSTYAEPPSRFEAGTPAIGEAIALGAAVDYLSGIG----MPKIHEYEVEIGKYLYEKLSS-LPDVRIYGPRP-----
PFALC    SGHKFCASLGTGFIFINKELSSKYKFKPLLYGSNIITNVSKYKSKFVTSLSELLETGTQNIPGILSMGISLEFFKKINWN-YVYQYEMYLYDLFIYYMNKYMK-NHFVQLPNLNLSYKKE
PKNOW    SGHKFCAPFGSGFVYIKKKITHASKTNPFLYGSNIITDVNKYRSKFVSS-PHIFETGTQNISAILAMGEAIKYLQKINDEGNAYKYEMYLYDLFLFYLRSFLT-HHLVELPNIPTKLD--
PYOEL    SGHKFCAPLGCGFIYIKNTLTCSYKFKPLLYGSNMITTVGKYKSEFVSS-PHLFETGTQNIAGVISMGVALDFLEKIDKN-LLCRYEMYLYDVLIYYLTKHLQ-RGLVQLPGG-------
TGOND    SGHKMYGPTGVGFLYGKYELLRNMP--PWKGGGEMIEFVDL-CESTYANPPARFEAGTPPFLQVIGLGAAVDFIEEVG----WPAIYSHDARLQRALHEVLASRFPELRLFCPPSAGTAY

ECOLI    ----------------------------------------------------------NRLGVIAFNLGKHHAYDVGSFLDNY-GIAVRTGHHCA
STYPH    ----------------------------------------------------------MRDRGSPVAFVVEGVHAHDVGQVLDDG-GVAVRVGHHCA
YPEST    ----------------------------------------------------------ERAGVIAFNLGQHHAYDVGSFLDQY-GIAIRTGHHCA
ECHRY    ----------------------------------------------------------QRLGVIAFNLGKHHAYDVGSFLDNY-GIAVRTGHHCA
MTUBE    ----------------------------------------------------------HRQGVIAFNLGRHHAFDVGSFLDQY-GIAIRTGHHCA
ATHAL    ----------------------------------------------------SESVHRGALCSFNVEGLHPTDLATFLDQQHGVAIRSGHHCA
PFALC    NINYKSHMQTHPPVHKYNDEQNFTNDHNITQSK---QTKSIHSQHDTFKIYTHDTRKYGLKKIGILPLWSNTFSSFDLVTFLDFK-NICIRAGHHCA
PKNOW    ------------EMHTSXQIRNNKVDDTYTHSE------IQNEDNDYLKIFVHNTRKDNKKKIGILPLWSLNFTSFDLVTFLDFK-NICIRSGHHCA
PYOEL    --------------VKG-DMMNEKEDTTLGRS----------SENDQCRLYIHNSRMVKGKKVPILPLWSDQFTSYDLVTFLDFK-NVCIRSGHHCA
TGOND    APELLSDAFDFSKVKQHRWVQSDQSDGRSAVTANAVHRSAVAGATDEARNGAFASGIERIPSIPLISFAHPEIHAHDIAVFLDVCGGVCVRSGHHCC
```

SOE-PCR (Fig. 5.10) was used to generate PfSufS[M]χ, with the PrP-like loop excised and replaced with a six amino acid bridge copied from the EcSufS sequence (intended to ease folding). Homology modelling suggested that the chimaeric protein would be able to adopt a native conformation. PfsufS[M]χ was cloned into several vectors for expression in *E. coli* and *S. cerevisiae*. Experiments with these vectors are now described. Schematic representations of PfSufS[M]χ fusion proteins are given in Fig. 5.11.

5.3.5. Expression and purification experiments with pET43.1b(+):PfsufS[M]χ.

5.3.5.1. Properties of the fusion protein.

Properties of the fusion partners in pET43.1b(+) constructs are discussed in section 5.3.3.3. NusA fusion again produced a significant increase in the Wilkinson-Harrison score of PfSufS[M]χ (18.5 → 55.5 %).

5.3.5.2. Protein expression.

Induction of cells at 37 °C led to the accumulation of protein with a functional S-tag (Fig. 5.12A). Image analysis of Coomassie-stained poly-acrylamide gels indicated the M_r to be 114 kDa; close to the predicted value of 111 kDa. However, the NusA-PfSufS[M]χ fusion apparently offered no solubilising benefit (Fig. 5.12B).

FIGURE 5.10.

CONSTRUCTION OF PFSUFS[M]χ BY SPLICING BY OVERLAP EXTENSION-PCR.

PfsufS[M]χ was assembled in three distinct steps by splicing by overlap extension (SOE)-PCR. In the first step, sequence encoding the N-terminal (N), middle (M) and C-terminal (C) of PfsufS[M] was recovered by RT-PCR (panel 1). A second step was required to assemble M and C regions with a short stretch of *E. coli* sequence (*E. coli* bridge), giving M+χ and C+χ, respectively (panel 2). Products of the first two PCR steps were assembled (As) in a final step, giving PfSufS[M]χ(panel 3). Additional abbreviations are as follows. bp: base pair; M: 123 bp ladder (Sigma).

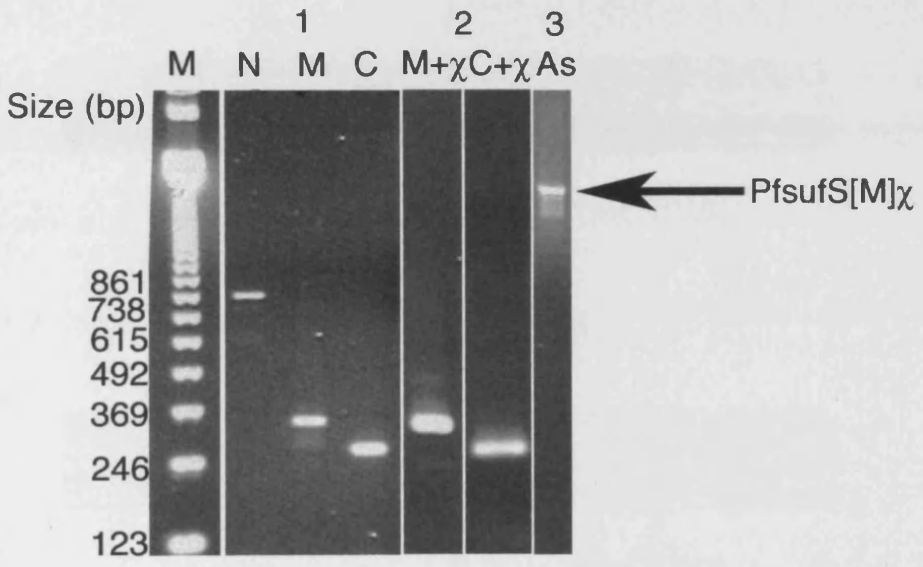

1
2
3
M
N
M
C
M+χ
C+χ
As
Size (bp)
861
738
615
492
369
246
123
PfsufS[M]χ

FIGURE 5.11.

SCHEMATICS OF PFSUFS[M]χ FUSION PROTEINS.

Thrombin, enterokinase, rhinovirus 3C protease ($3C^{pro}$) and Factor Xa sites are indicated. Abbreviations are as follows. NusA: N-utilising substance A; GST: glutathione *S*-transferase; cMBP: cytosolic maltose binding protein.

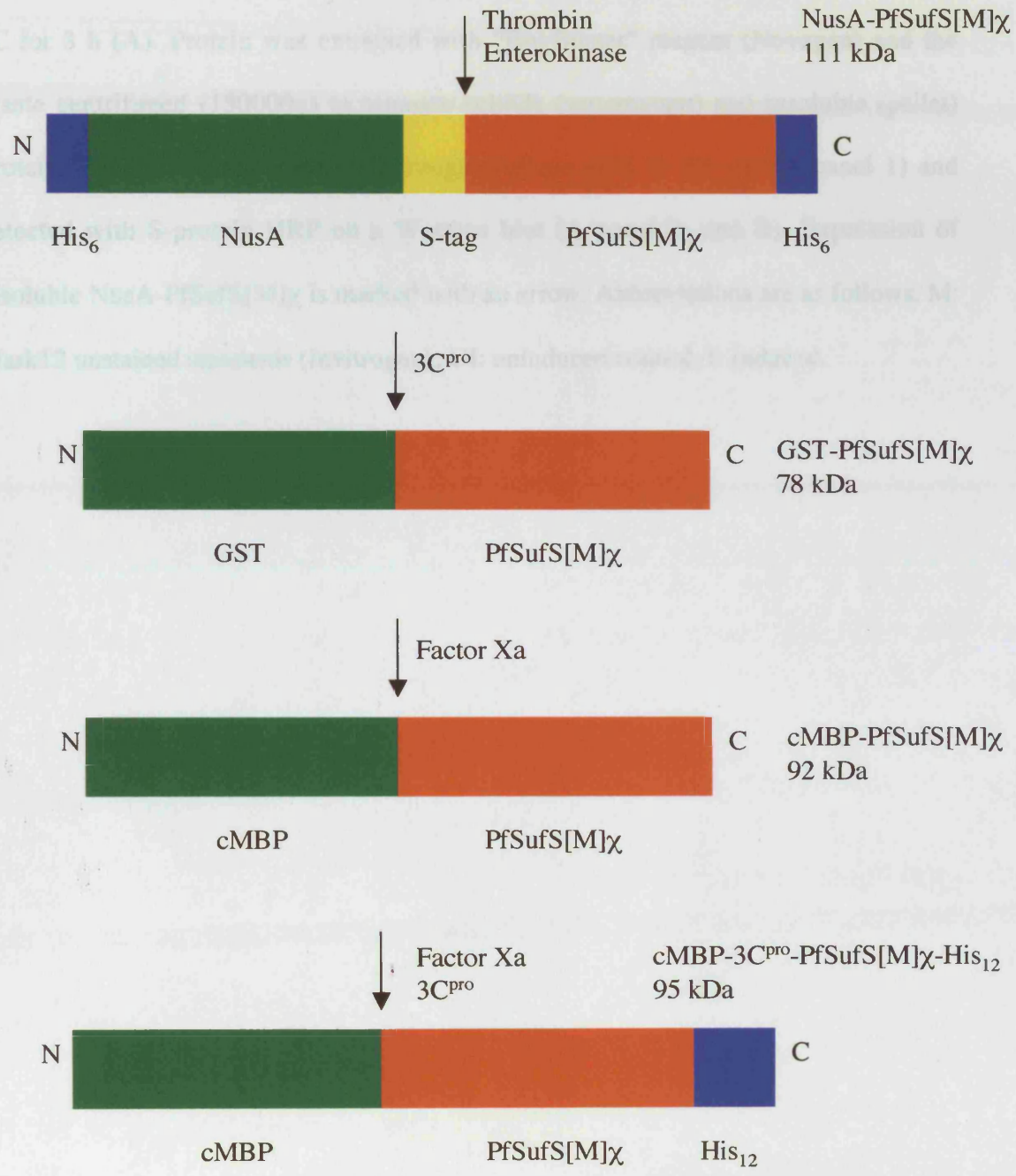
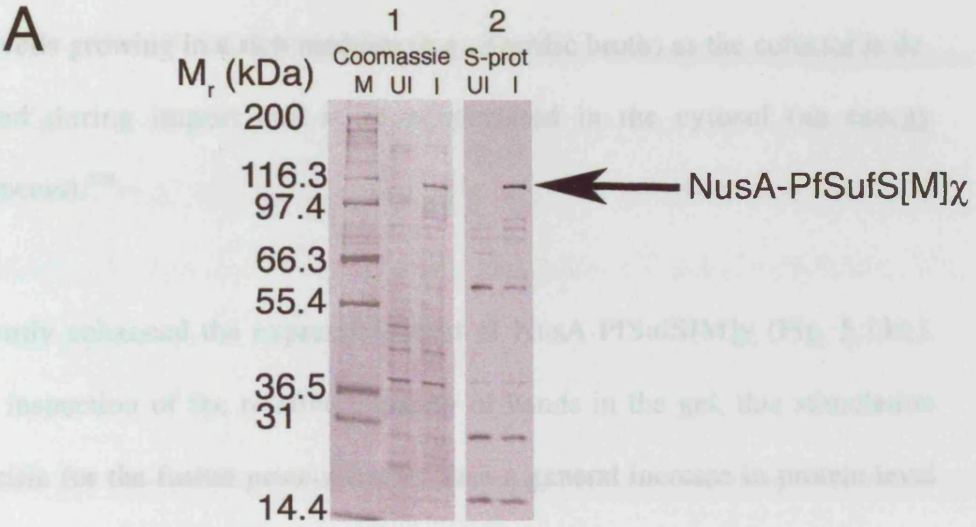
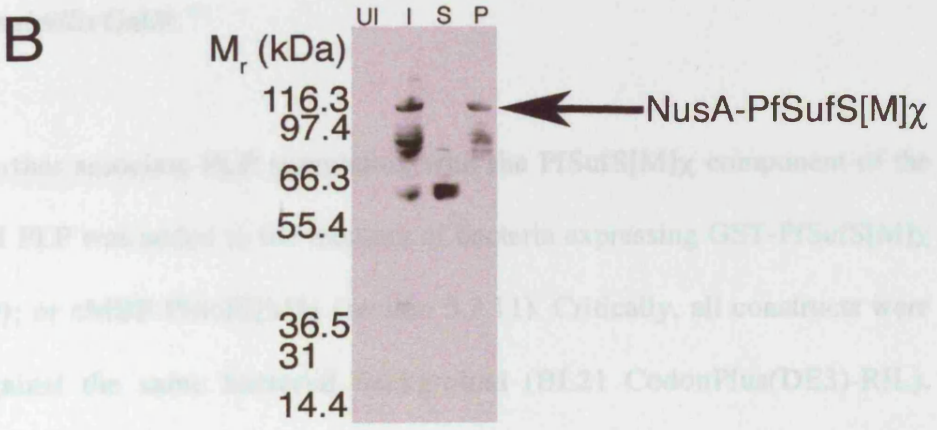

Thrombin
Enterokinase
NusA-PfSufS[M]χ
111 kDa
N
C
His6
NusA
S-tag
PfSufS[M]χ
His6
3Cpro
N
C
GST-PfSufS[M]χ
78 kDa
GST
PfSufS[M]χ
Factor Xa
N
C
cMBP-PfSufS[M]χ
92 kDa
cMBP
PfSufS[M]χ
Factor Xa
3Cpro
cMBP-3Cpro-PfSufS[M]χ-His12
95 kDa
N
C
cMBP
PfSufS[M]χ
His12

FIGURE 5.12.

EXPRESSION OF NUSA-PFSUFS[M]χ.

E. coli cells harbouring pET43.1b(+):PfsufS[M]χ were induced with 1 mM IPTG at 37 °C for 3 h (A). Protein was extracted with "BugBuster" reagent (Novagen) and the lysate centrifuged (150000*g*) to separate soluble (supernatant) and insoluble (pellet) protein (B). Protein was resolved through NuPage 4-12 % BT gel (A panel 1) and detected with S-protein HRP on a Western blot (A panel 2; and B). Expression of insoluble NusA-PfSufS[M]χ is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); UI: uninduced control; I: induced.


A
1
2
Coomassie
S-prot
M_r (kDa)
M
UI
I
UI
I
200
116.3
97.4
66.3
55.4
36.5
31
14.4
NusA-PfSufS[M]χ


B
UI
I
S
P
M_r (kDa)
116.3
97.4
66.3
55.4
36.5
31
14.4
NusA-PfSufS[M]χ

5.3.5.3. Specific stimulation of protein expression by PLP.

To enhance expression, I enriched the bacterial growth medium with 50 μM PLP.* PLP was chosen for trivial reasons of availability; any of the B_6 vitamers could be satisfactorily absorbed by bacteria.[370] Notably in vitamin B_6 auxotrophs, PLP is only acceptable to cells growing in a rich medium (e.g., Terrific broth) as the cofactor is de-phosphorylated during import and re-phosphorylated in the cytosol (an energy demanding process).[370]

PLP significantly enhanced the expression level of NusA-PfSufS[M]χ (Fig. 5.13A). Indeed, from inspection of the relative intensity of bands in the gel, this stimulation appeared specific for the fusion protein (rather than a general increase in protein level due to faster bacterial growth). Consideration of the nature of the fusion protein indicated that PLP could be interacting with NusA or PfSufS[M]χ. While, to my knowledge, there is no published evidence suggesting that PLP interacts with NusA, it should be understood that several prokaryotic transcription factors are PLP-dependent, e.g., *Bacillus subtillis* GabR.[371]

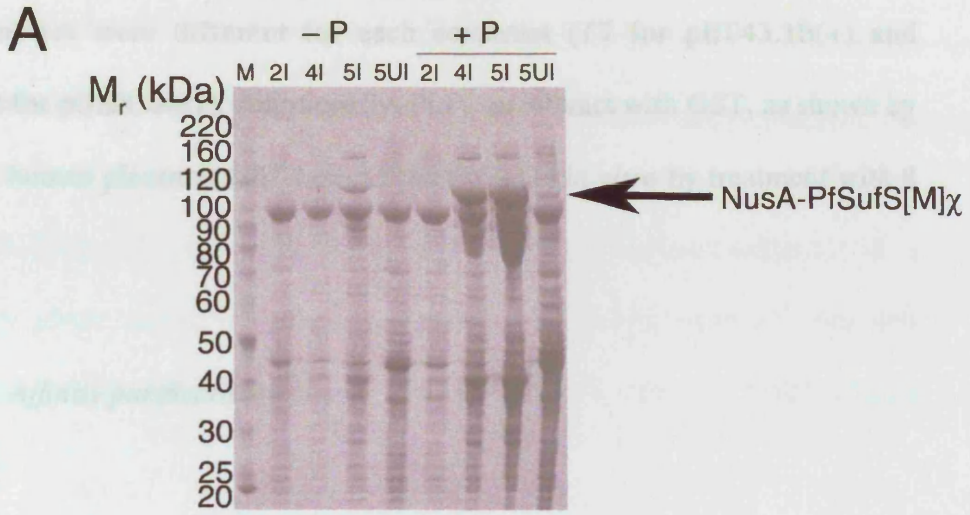
In order to further associate PLP stimulation with the PfSufS[M]χ component of the fusion, 50 μM PLP was added to the medium of bacteria expressing GST-PfSufS[M]χ (section 5.3.9); or cMBP-PfSufS[M]χ (section 5.3.11). Critically, all constructs were expressed against the same bacterial background (BL21 CodonPlus(DE3)-RIL). Limited enhancement of expression was observed for GST-PfSufS[M]χ and cMBP-PfSufS[M]χ (Fig. 5.13B).

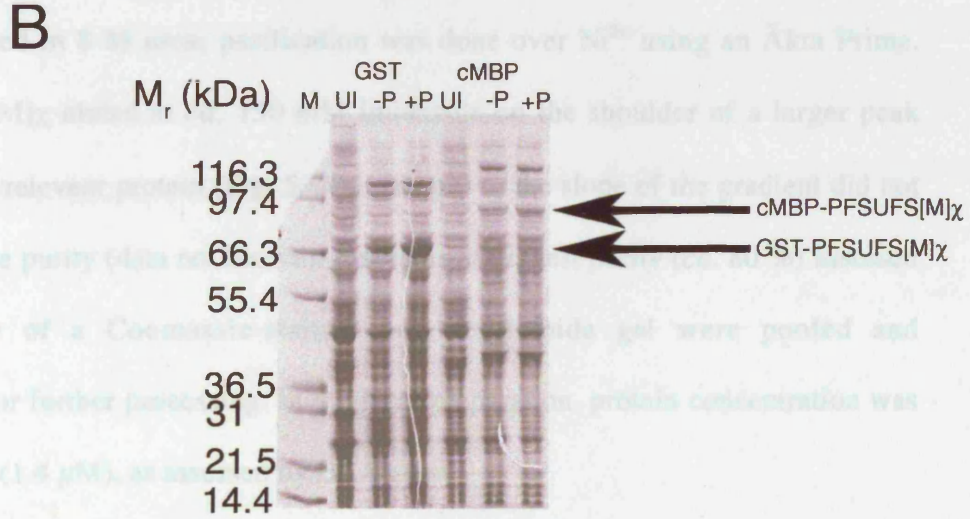
* PLP addition is advocated by Dr. R. Lill (Philipps-Universität Marburg, Germany) and (the late) Dr. G. Kispal.

FIGURE 5.13.

PFSUFS[M]χ EXPRESSION WITH PYRIDOXAL 5'-PHOSPHATE.

E. coli cells harbouring pET43.1b(+):PfsufS[M]χ, pGEX-6-P1:PfsufS[M]χ and pMALc2:PfsufS[M]χ were induced with 1 mM IPTG at 37 °C for 3-5 h in medium enriched with 50 μM PLP. Protein was resolved through NuPage 4-12 % BT gel. Expression of NusA-PfSufS[M]χ (A), GST-PfSufS[M]χ (B) and cMBP-PfSufS[M]χ (B) was enhanced by PLP. Abbreviations are as follows. M: SeeBlue (A) or Mark12 unstained standards (B); 2, 4, 5: 2, 4, and 5 h post-induction; UI: uninduced control; I: induced; –P: PLP absent; +P: 50 μM PLP.


A
– P
+ P
M 2I 4I 5I 5UI 2I 4I 5I 5UI
M$_r$ (kDa)
220
160
120
100
90
80
70
60
50
40
30
25
20
NusA-PfSufS[M]χ


B
GST
cMBP
M$_r$ (kDa)
M UI -P +P UI -P +P
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PFSUFS[M]χ
GST-PFSUFS[M]χ

Therefore, the possibility exists that PLP-stimulation is specific to PfSufS[M]χ, independent of fusion partner. Therefore, I speculate that PLP is stabilising folding intermediates of PfSufS[M]χ, thus abrogating degradation. Hence, PfSufS[M]χ fusion proteins could be binding PLP in *E. coli* cultures. Several caveats need to be inserted. Notably, promoters were different for each construct (T7 for pET43.1b(+) and pMALc2; *Ptac* for pGEX-6-P1). Additionally, PLP can interact with GST, as shown by inactivation of human placental GSTπ (GenBank P09211) *in vitro* by treatment with 8 mM PLP.[372]

5.3.5.4. Affinity purification.

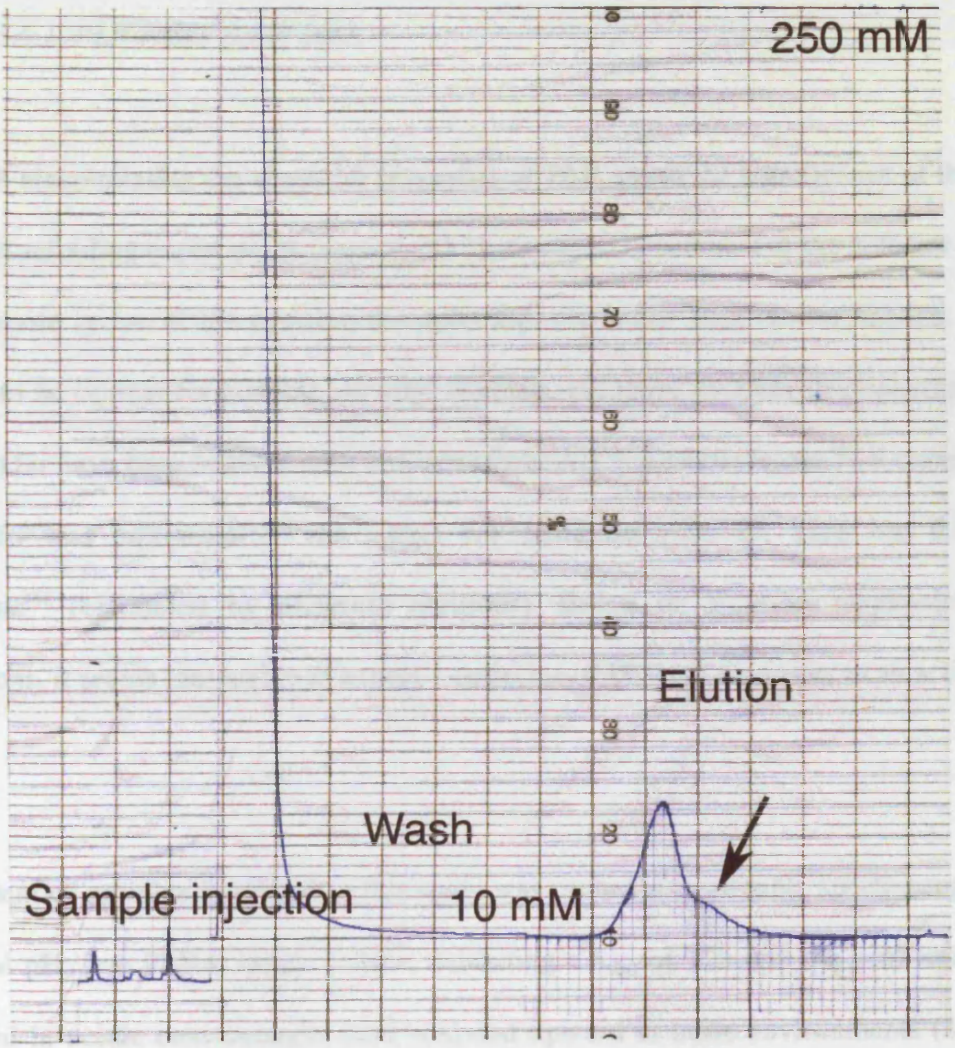
In a typical preparative purification of NusA-PfSufS[M]χ, cells were broken via passage through a French pressure cell. However, *ca.* 10 L of culture (*ca.* 50 g wet weight of bacteria) was required to give sufficient protein for three experiments; this highlights the typically low yield achieved with heterologously expressed *P. falciparum* proteins. Inclusion bodies were recovered by repeated washing and centrifugation and finally dissolved in 8 M urea; purification was done over Ni^{2+} using an Äkta Prime. NusA-PfSufS[M]χ eluted at *ca.* 150 mM imidazole on the shoulder of a larger peak composed of irrelevant protein (Fig. 5.14). Changes to the slope of the gradient did not greatly improve purity (data not shown). Fractions of highest purity (*ca.* 80 %) assessed by inspection of a Coomassie-stained poly-acrylamide gel were pooled and concentrated for further processing. In a typical preparation, protein concentration was ~ 150 $\mu g.mL^{-1}$ (1.4 μM), as assessed by BCA assay.

FIGURE 5.14.

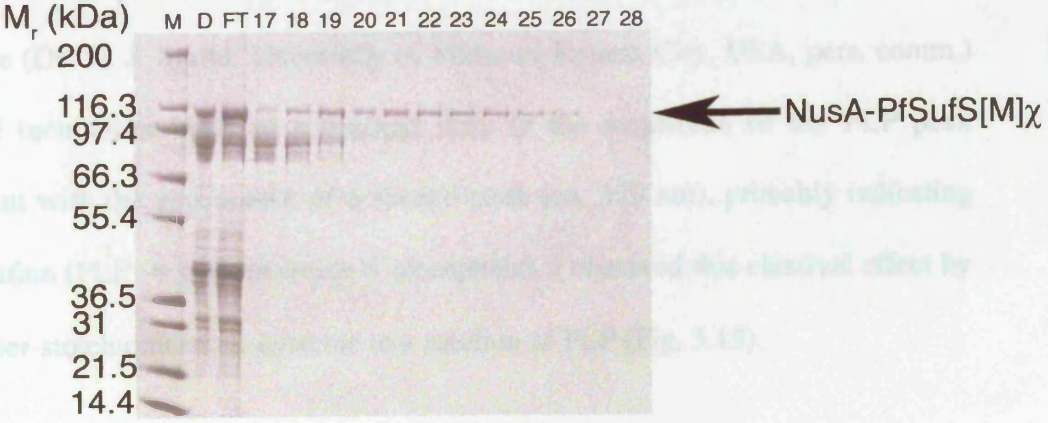
PURIFICATION OF INSOLUBLE NUSA-PFSUFS[M]χ.

E. coli cells harbouring pET43.1b(+):PfsufS[M]χ were induced with 1 mM IPTG for 5 h at 37 °C. Inclusion bodies were dissolved with 8 M urea. Insoluble NusA-PfSufS[M]χ was purified over Ni^{2+} by FPLC with a linear gradient limited by 250 mM imidazole. A typical 280 nm elution profile is shown (A). Elution fractions were resolved through NuPage 4-12 % BT gel (B). Evidence of the fusion protein in elution fractions > 18 is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); D: dissolved inclusion bodies; FT: column flow-through; 17-28: elution fractions 17-28.

A


250 mM
Elution
Wash
Sample injection
10 mM

B


M D FT 17 18 19 20 21 22 23 24 25 26 27 28
M_r (kDa)
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
NusA-PfSufS[M]χ

5.3.6. Pyridoxal 5′-phosphate.

5.3.6.1. PLP-lysine Schiff base.

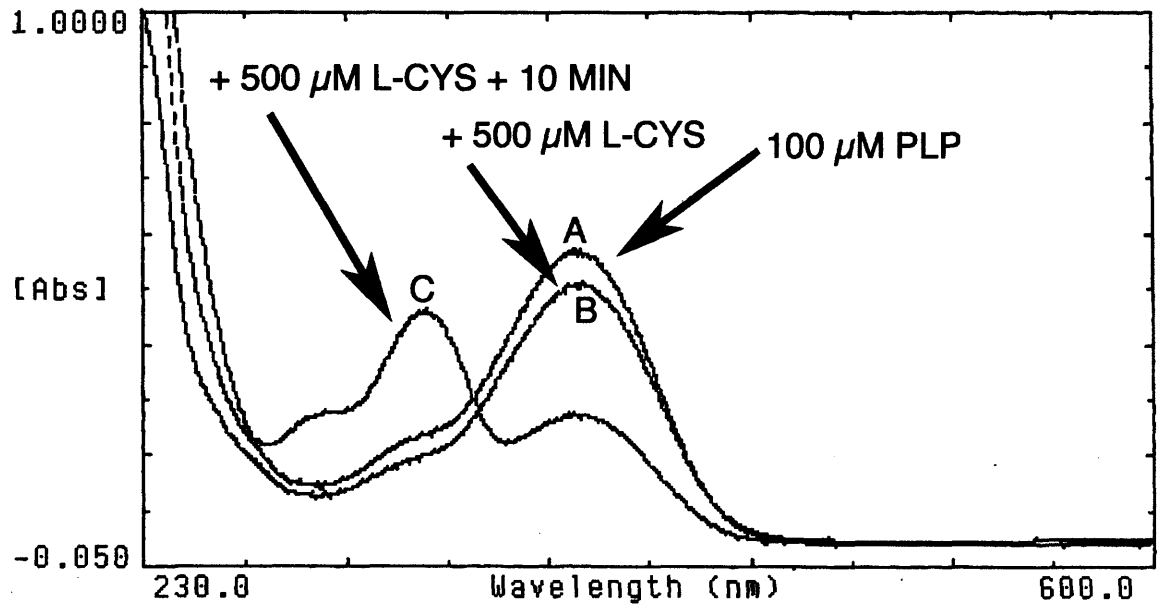
It is worth recapitulating the chemical properties of PLP, given the significance of this cofactor in refolding experiments (section 5.3.7). Covalent attachment of the cofactor to the active site lysine of the cysteine desulphurase, via Schiff base (C=N) linkage, is assumed to be critical for enzyme activity. Notably, all native cysteine desulphurases studied so far have been isolated with PLP already bound. For example, the PLP content of EcSufS was equimolar;[165] the value for *Synechocystis* sp. SufS was 0.58 equivalents[311] (estimated by different methods). Based on *a priori* mechanistic assumptions, it seems reasonable to expect 1 molecule of PLP per monomer in an active enzyme.

PLP content of proteins described in this study was probed using visible spectroscopy. At neutral pH, free PLP is bright yellow, exhibiting a λ_{max} of 325 nm; the peak shifts downward in acidic environments (295 nm) and upward in basic environments (388 nm).[373] Another type of absorbance spectrum develops when free PLP reacts with primary amines (e.g., amino acids), forming Schiff bases. Specifically, λ_{max} shifts to *ca.* 415 nm (neutral pH).[374] As such, a λ_{max} of *ca.* 415 nm may be considered diagnostic of a Schiff base (Dr. A. J. Iriarte, University of Missouri-Kansas City, USA, pers. comm.) Prolonged incubation leads to a gradual drop in the amplitude of the PLP peak concomitant with the appearance of a second peak (*ca.* 330 nm), probably indicating transamination (PLP → pyridoxamine 5′-phosphate). I observed this classical effect by adding super-stoichiometric L-cysteine to a solution of PLP (Fig. 5.15).

FIGURE 5.15.

ABIOTIC FORMATION OF SCHIFF BASE BETWEEN PYRIDOXAL 5′-PHOSPHATE AND L-CYSTEINE.

PLP (100 μM) in aqueous buffer (PBS) was mixed with super-stoichiometric L-cysteine (500 μM) at 25 °C in a quartz-glass cuvette; absorbance spectra were acquired. Instantaneous transition from a single peak of free PLP (A) to a red-shifted peak of lower amplitude (Schiff base, B) upon addition of L-cysteine (L-cys) is observed. A pyridoxamine 5′-phosphate (PMP) peak appears in the lower energy regime of the spectrum after 10 min incubation (C).


1.0000
[Abs]
-0.050
230.0
Wavelength (nm)
600.0
+ 500 μM L-CYS + 10 MIN
+ 500 μM L-CYS
100 μM PLP
A
B
C

Formation of a Schiff base by condensation of an aldehyde (PLP) with a primary amine does not require aggressive conditions.[375] Indeed, Schiff base linkages are readily formed with amino acids, most stably with lysine (by attack upon the ε group). However, reaction of PLP with polypeptides is a more challenging reaction. PLP is inert towards polypeptides with the important exception of random coils of poly-lysine.[376] Hence, in refolding experiments, adventitious binding to amino acids other than lysine may be excluded on chemical grounds. However, NusA-PfSufS[M]χ possesses 78 lysine residues. Any lysines exposed to solvent might be accessible to electrophilic attack by PLP. However, the Schiff base formed with the catalytic lysine probably exhibits greatest stability because the phosphate group of the PLP makes additional hydrogen bonding contacts with the active site cleft.

5.3.6.2. PLP in protein refolding.

It is unknown if PLP binding is an essential feature of the folding pathway of SufS. Classical studies have been conducted on the role of PLP in the *in vitro* folding of enzymes of the aminotransferase class, notably, aspartate aminotransferase (AAT; EC 2.6.1.1). For example, Reyes, *et al.*,[377] found that a greater amount of active rat mitochondrial AAT was recovered after refolding in the presence than in the absence of PLP; DTT was also vital.

However, it was evident that PLP binding was not absolutely necessary for acquisition of a native fold. Indeed, yields of functional *E. coli* AAT were high even if PLP was absent from the refolding buffer.[378] Nevertheless, once the native fold is achieved, PLP is probably an important determinant of structural stability.[379]

5.3.7. Refolding experiments with NusA-PfSufS[M]χ.

5.3.7.1. Refolding screen.

Purified NusA-PfSufS[M]χ was subjected to a commercially available factorial refolding screen; 8 buffers were randomly supplemented with PLP. Protein solubility was assessed by partition into the supernatant after centrifugation at 150000*g*, as indicated by Western blot (Fig. 5.16). Six of the 16 refolding buffers (4, 6, 8, 10, 12 and 16) solubilised NusA-PfSufS[M]χ. However, no common factor was apparent.

5.3.7.2. Preparative refolding.

The first four solubilising buffers (4, 6, 8 and 10) were arbitrarily chosen for initial scale-up experiments, with the object of testing refolded protein for enzymatic activity. Preparative refolding (Fig. 5.17) gave qualitatively similar results to the refolding screen (section 5.3.7.1). After extensive concentration of refolded soluble protein, typical concentrations estimated by absorbance at 280 nm were 7-14 μM (*ca.* 0.8-1.5 mg.mL^{-1}). PLP was added to a high concentration (0.5 mM) in refolding buffers to mitigate any effects of dephosphorylation.

FIGURE 5.16.

REFOLDING SCREEN OF NUSA-PFSUFS[M]χ.

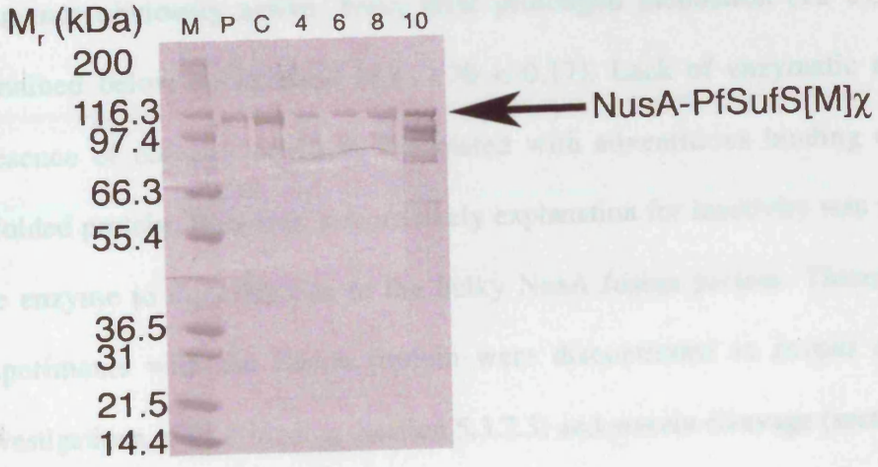
Insoluble NusA-PfSufS[M]χ dissolved in 8 M urea (Fig. 5.14) was subjected to a factorial refolding screen comprising 16 different buffers (compositions stated in Table 5.2). Refolded protein was centrifuged (150000*g*) to fractionate dissolved (supernatant) and precipitated (pellet) protein. Dissolved protein was resolved through NuPage 4-12 % BT gel and detected by Western blot using S-protein HRP conjugate (Novagen) Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); 1-16: soluble protein refolded in buffers 1-16.


M_r (kDa)
M 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
NusA-PfSufS[M]χ

FIGURE 5.17.

PREPARATIVE REFOLDING OF NUSA-PFSUFS[M]χ.

Pooled fractions of insoluble NusA-PfSufS[M]χ dissolved in 8 M urea (Fig. 5.14) were concentrated and refolded by dialysis against buffer 4, 6, 8 and 10 (compositions stated in Table 5.2). Refolded soluble protein was resolved through NuPage 4-12 % BT gel. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); P: pooled fractions of insoluble NusA-PfSufS[M]χ; C: 5 × concentrated P; 4, 6, 8 and 10: soluble protein refolded in buffers 4, 6, 8 and 10.


M_r (kDa)
M P C 4 6 8 10
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
NusA-PfSufS[M]χ

Protein refolded in FoldIt buffer 10 exhibited a λ_{max}cofactor (395 nm) suggestive of free PLP (Fig. 5.18D). Conversely, the higher values of λ_{max}cofactor for protein refolded in FoldIt buffers 4 (416 nm), 6 (410 nm) and 8 (416 nm) were robustly indicative of Schiff base rather than free PLP (Fig. 5.18ABC). Slight differences probably reflect cofactor protonation: buffers 4 and 8 were held at pH 8.2; while buffer 6 was held at pH 6.5. Mihara, *et al.*[313] reported little significant difference in the λ_{max}cofactor of EcSufS in the range pH 6-8. Given some evidence for the presence of PLP Schiff base, NusA-PfSufS[M]χ refolded in buffer 4, 6 and 8 was tested for enzyme activity by MB^+ assay. Protein was diluted 1/10 into the assay, giving a concentration of *ca.* 700 nM (matches protein concentration used in control assays in chapter 4). However, refolded proteins were not obviously active. Even after prolonged incubation (12 h), MB^+ signals remained below background (Abs 670 < 0.17). Lack of enzymatic activity in the presence of cofactor could be associated with adventitious binding of PLP to the refolded protein. However, a more likely explanation for inactivity was the inability of the enzyme to dimerize due to the bulky NusA fusion partner. Therefore, refolding experiments with the fusion protein were discontinued in favour of preliminary investigations of PLP binding (section 5.3.7.3) and protein cleavage (section 5.3.8).

5.3.7.3. Adventitious PLP binding to NusA-PfSufS[M]χ.

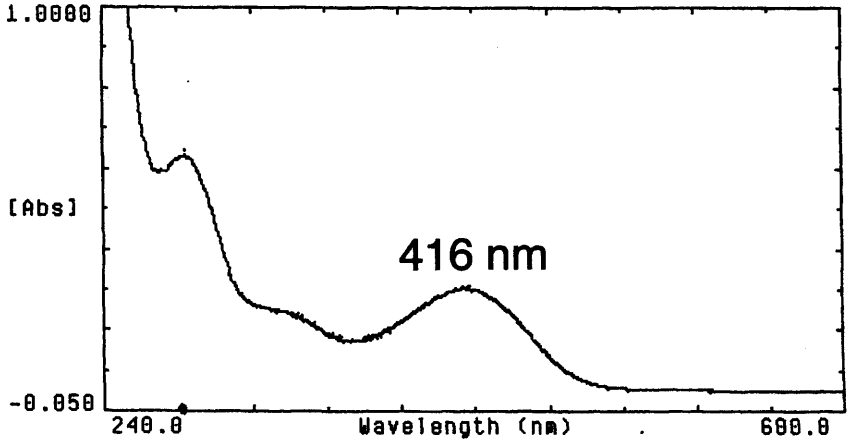
In order to investigate further the phenomenon of PLP interaction with NusA-PfSufS[M]χ, a crude estimate of PLP binding, the ratio of λ_{max}280 to λ_{max}cofactor, was taken. Ratios obtained for FoldIt 4 (0.44), 6 (0.27) and 8 (0.83) were variable; no correlation between the two values was evident.

FIGURE 5.18.

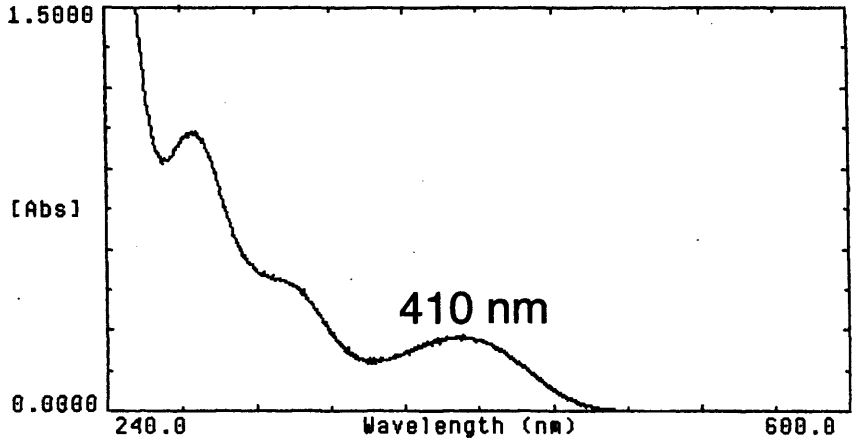
ABSORBANCE SPECTRA OF REFOLDED NUSA-PFSUFS[M]χ.

Absorbance spectra of soluble NusA-PfSufS[M]χ (Fig. 5.17) refolded in buffer 4 (A), 6 (B), 8 (C) and 10 (D) were acquired. Schiff base peaks are apparent from protein refolded in buffers 4, 6 and 8. Conversely, the peak at 395 nm exhibited by protein refolded in buffer 10 suggests free PLP.

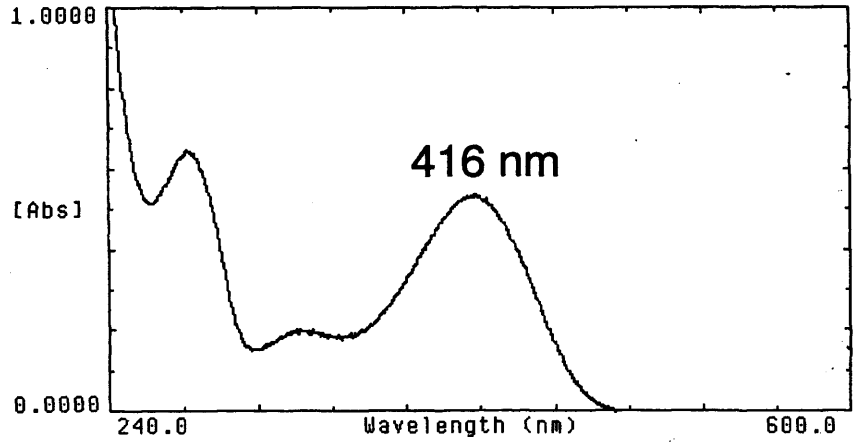
A


1.0000
[Abs]
-0.050
416 nm
240.0
Wavelength (nm)
600.0

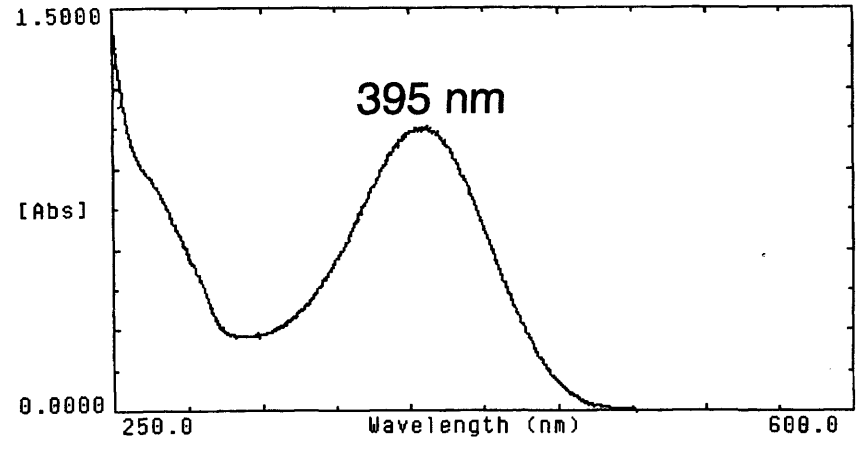
B


1.5000
[Abs]
0.0000
410 nm
240.0
Wavelength (nm)
600.0

C


1.0000
[Abs]
0.0000
416 nm
240.0
Wavelength (nm)
600.0

D


1.5000
[Abs]
0.0000
395 nm
250.0
Wavelength (nm)
600.0

Therefore, it seemed that binding efficiency varied between refolding conditions; or that adventitious (non-stoichiometric) binding was occurring. Indeed, it has recently been noted that dialysis is a poor method for removal of adventitiously bound PLP and that gel filtration is preferable (Dr. A. J. Iriarte, pers. comm.).

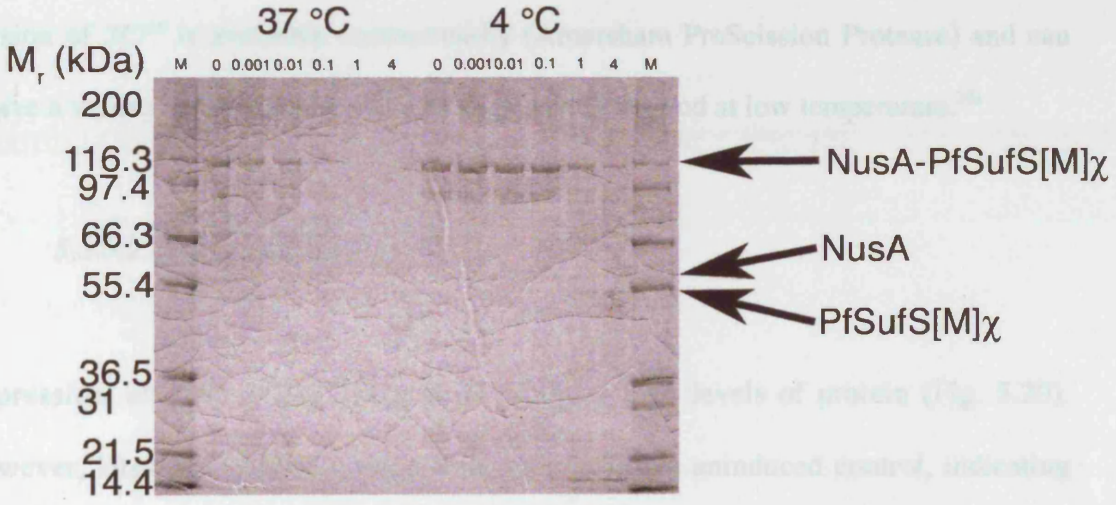
5.3.8. Cleavage experiments with NusA-PfSufS[M]χ.

Given the possibility that the large NusA tag was interfering with dimerization or folding, attempts were made to cleave NusA from PfSufS[M]χ via digestion with EnterokinaseMax (Fig. 5.19). To this end, protein dissolved in 8 M urea was partially re-natured by dialysis for 40 h against buffer containing 1 M urea. Following dialysis, aliquots of protein were digested for 20 h under various conditions, namely, high (37 °C) and low (4 °C) temperature; and differing amounts of EnterokinaseMax (0.001-4 U per reaction). Inefficient cleavage of the fusion protein into 58 kDa and 55 kDa fragments was detectable following digestion at 4 °C with the largest amount of protease (4 U). Conversely, protein was completely degraded by 1 U EnterokinaseMax at 37 °C. Taken together, these data indicate that NusA-PfSufS[M]χ can be cleaved. Conversely, in the absence of protease, the fusion protein was stable for long periods at 37 °C (data not shown).

FIGURE 5.19.

ENTEROKINASE CLEAVAGE OF NUSA-PFSUFS[M]χ.

Purified NusA-PfSufS[M]χ (Fig. 5.14) was cleaved by incubation with enterokinase at 37 °C and 4 °C. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); 0, 0.001, 0.01, 0.1, 1 and 4: 0, 0.001, 0.01, 0.1, 1 and 4 U enterokinase.


37 °C
4 °C
M_r (kDa)
M 0 0.001 0.01 0.1 1 4 0 0.001 0.01 0.1 1 4 M
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
NusA-PfSufS[M]χ
NusA
PfSufS[M]χ

5.3.9. Experiments with pGEX-6P-1:PfsufS[M]χ.

5.3.9.1. Properties of the fusion protein.

The pGEX-6P-1 vector allows the expression of PfSufS[M]χ N-terminal fused to the 26 kDa C-terminus of *Schistosoma japonicum* glutathione *S*-transferase (GST; GenBank AAB37346).[380] GST can be used to purify the fusion by specific interaction with glutathione immobilised to a resin. The two fusion partners are separated by an 11-amino acid linker containing a recognition site for the protease $3C^{pro}$. Hence, the GST tag may be cleaved from PfSufS[M]χ by specific proteolytic digestion. A GST-tagged version of $3C^{pro}$ is available commercially (Amersham PreScission Protease) and can cleave a variety of fusion proteins with high specificity and at low temperature.[381]

5.3.9.2. Protein expression.

Expression of GST-PfSufS[M]χ at 37 °C gave high levels of protein (Fig. 5.20). However, large amounts of protein were present in the uninduced control, indicating that expression was unregulated. As determined by image analysis of Coomassie-stained gels, the M_r of the protein was 71 kDa. This approximates the predicted value (78 kDa).

Soluble protein was recovered from cells by extraction in detergents and subjected to immunoblot with anti-GST antibody (Fig. 5.21). Cross-reactivity was demonstrated, suggesting that GST-PfSufS[M]χ was expressed.

FIGURE 5.20.

EXPRESSION OF GST-PFSUFS[M]χ.

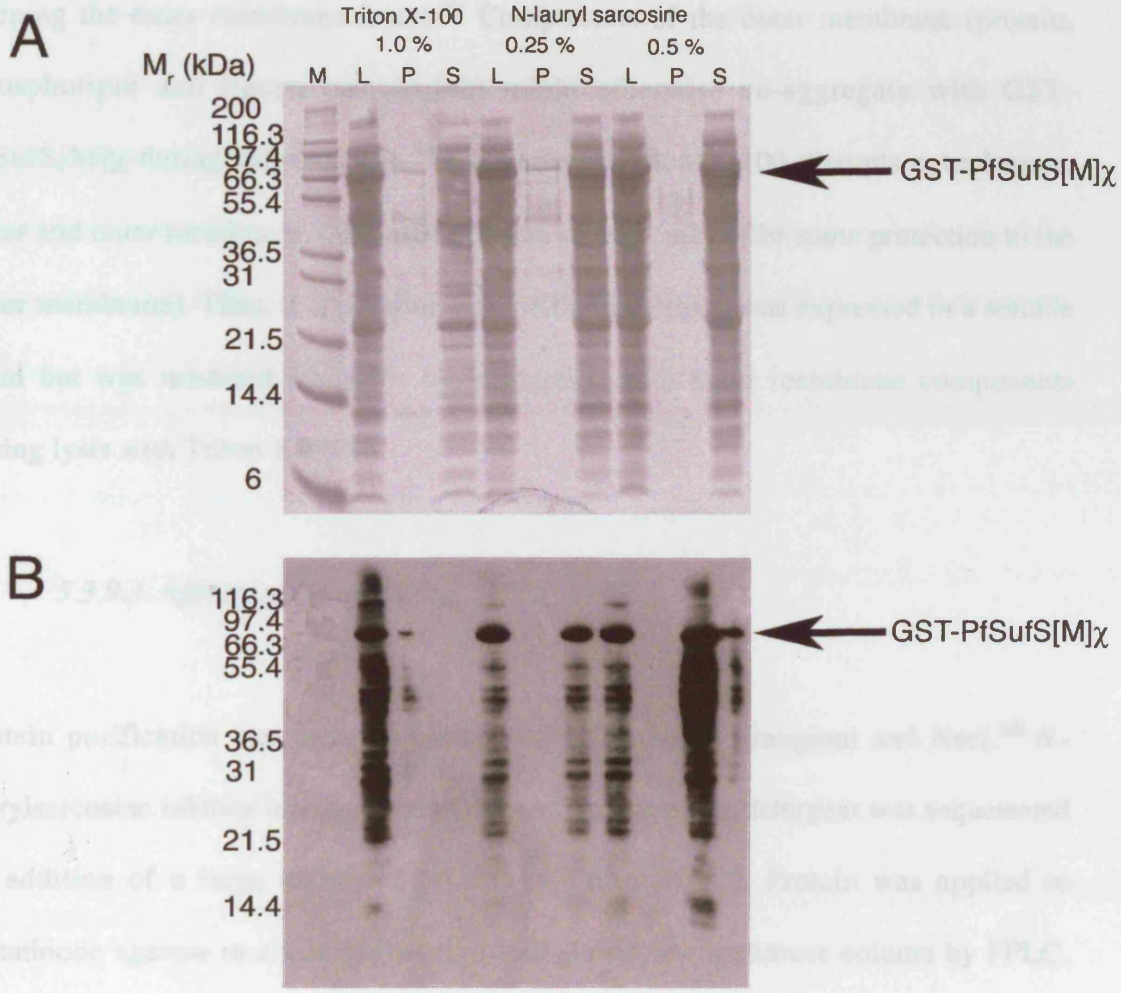
E. coli cells harbouring pGEX-6P-1:PfsufS[M]χ were induced with 1 mM IPTG at 37 °C. Samples were taken at 3, 4, and 5 h post-induction. Additionally, an uninduced culture was grown and a sample taken after 5 h of growth. Protein was resolved through NuPage 12 % BT gel. Evidence for expression of fusion protein (GST-PfSufS[M]χ) is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); 3, 4 and 5: samples taken 3, 4 and 5 h post-induction; UI: uninduced sample taken at 5 h.


M_r (kDa)
M
3 h
4 h
5 h
UI
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
GST-PfSufS[M]χ

FIGURE 5.21.

DETERGENT SCREEN OF GST-PFSUFS[M]χ.

E. coli cells harbouring pGEX-6P-1:PfsufS[M]χ were induced with 1 mM IPTG at 37 °C for 5 h. Protein was extracted in Triton X-100 (1.0 % v/v) or N-lauryl sarcosine (0.25 and 0.5 % w/v). Soluble (supernatant) protein was fractionated from insoluble (pellet) protein by centrifugation (150000*g*) and resolved through NuPage 12 % BT gel (A). Western blotting was done with anti-GST (B). B is a replicate blot of A. Evidence for expression of fusion protein is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); L: total lysate; P: pellet; S: supernatant.


A
Triton X-100
1.0 %
N-lauryl sarcosine
0.25 %
0.5 %
M r (kDa)
M L P S L P S L P S
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
GST-PfSufS[M]χ
B
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
GST-PfSufS[M]χ

However, a large amount of truncated protein was present as indicated by long smear of immunoreactivity. Additionally, I noted that protein was solubilised in ≥ 0.25 % w/v N-laurylsarcosine but not in 1 % v/v Triton X-100. Although N-laurylsarcosine is an anionic detergent, it is non-denaturing:[382] hence enzymatic activity may be retained.[383] The solubilising properties of N-laurylsarcosine can be attributed to the ability of the detergent to disrupt only the cytoplasmic and inner membranes of the bacterium while keeping the outer membrane intact.[384] Components of the outer membrane (protein, phospholipid and lipopolysaccharide) might otherwise co-aggregate with GST-PfSufS[M]χ during cellular lysis.[382] Conversely, Triton X-100 disrupts cytoplasmic, inner and outer membranes (although addition of Mg^{2+} may offer some protection to the outer membrane). Thus, it is possible that GST-PfSufS[M]χ was expressed in a soluble form but was rendered insoluble by interaction with outer membrane components during lysis with Triton X-100.

5.3.9.3. Affinity purification.

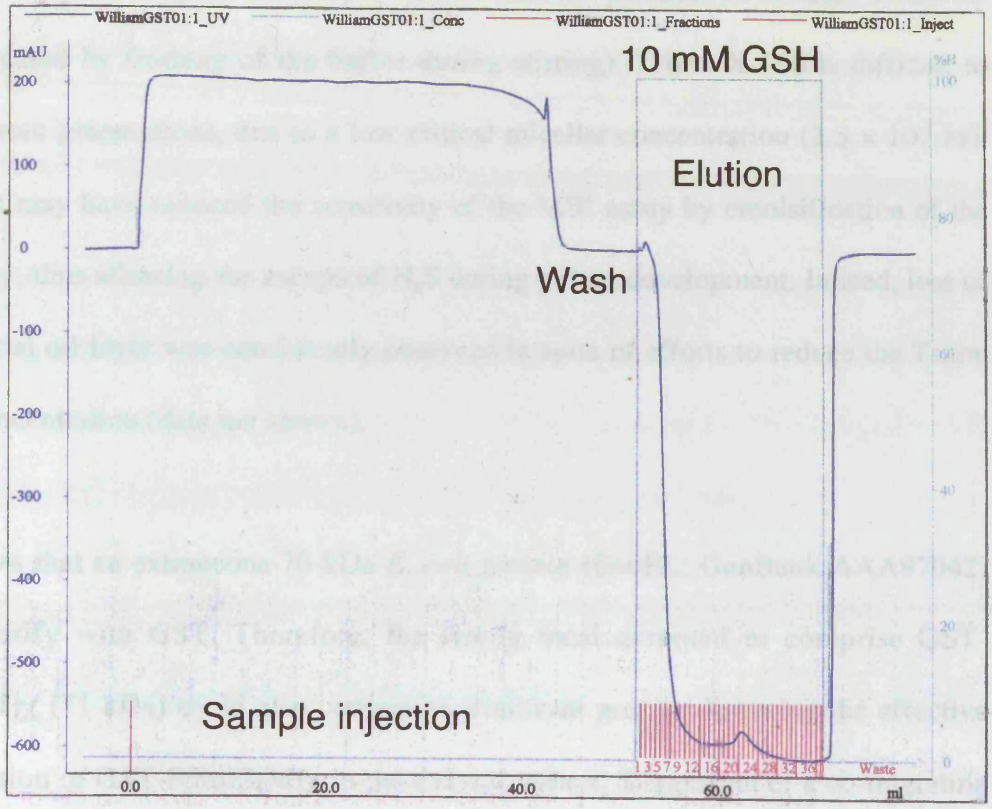
Protein purification was done essentially as indicated by Frangioni and Neel.[383] *N*-laurylsarcosine inhibits binding to glutathione. Therefore, the detergent was sequestered by addition of a large excess (2 % v/v) of Triton X-100. Protein was applied to glutathione agarose in a batch format; or to a glutathione sepharose column by FPLC. Affinity resin was washed in buffer containing Triton X-100 prior to elution with reduced glutathione. A typical FPLC run is summarised in Fig. 5.22. Baseline deflection is due to Triton X-100, which exhibits a λ_{max} of 275 nm.

FIGURE 5.22.

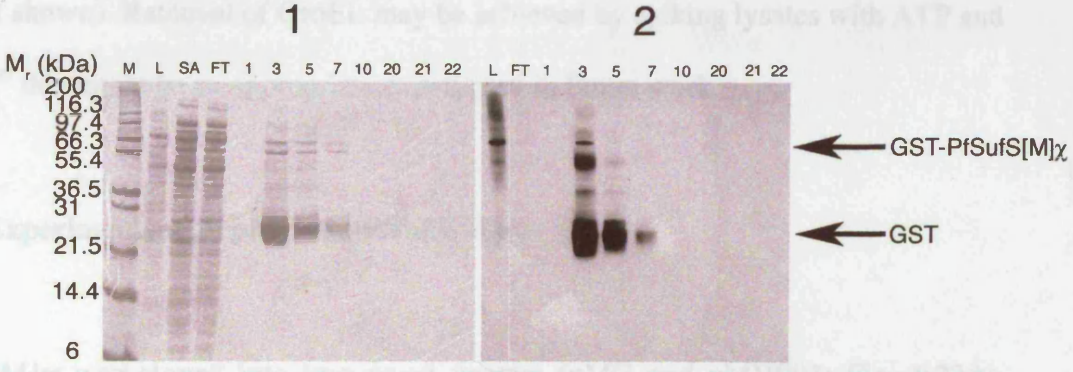
PURIFICATION OF GST-PFSUFS[M]χ.

E. coli cells harbouring pGEX-6P-1:PfsufS[M]χ were induced with 1 mM IPTG at 37 °C for 5 h. Protein was extracted in N-lauryl sarcosine (0.5 % w/v). Soluble protein was purified through a GSTrap column (Amersham) by FPLC and eluted with 10 mM reduced glutathione (GSH). A typical 280 nm elution profile is shown (A). Elution fractions were resolved through NuPage 4-12 % BT gel (B panel 1). Western blotting was done with anti-GST (B panel 2). Evidence of the fusion protein in elution fraction 3 is presented. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); L: lysed bacteria; FT: column flow-through; 1, 3, 5, 7, 10, 20, 21 and 22: elution fractions 1, 3, 5, 7, 10, 20, 21 and 22.

A


WilliamGST01:1_UV
WilliamGST01:1_Conc
WilliamGST01:1_Fractions
WilliamGST01:1_Inject
mAU
10 mM GSH
Elution
Wash
Sample injection
Waste
ml

B


1
2
M_r (kDa)
M L SA FT 1 3 5 7 10 20 21 22
L FT 1 3 5 7 10 20 21 22
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
6
GST-PfSufS[M]χ
GST

5.3.9.4. Enzymology.

Preparations were submitted to MB^+ assay. However, no signal was detected, indicating the absence of active cysteine desulphurase. I noted the presence of residual Triton X-100 (indicated by frothing of the buffer during stirring). Triton X-100 is difficult to remove from preparations, due to a low critical micellar concentration (2.5×10^{-4} M). Detergent may have reduced the sensitivity of the MB^+ assay by emulsification of the oil overlay, thus allowing the escape of H_2S during colour development. Indeed, loss of the essential oil layer was consistently observed in spite of efforts to reduce the Triton X-100 concentration (data not shown).

It is known that an extraneous 70-kDa *E. coli* protein (GroEL: GenBank AAA97042) can co-purify with GST. Therefore, the strong band assumed to comprise GST-PfSufS[M]χ (71 kDa) could also contain an irrelevant protein, lowering the effective concentration of GST-PfSufS[M]χ in the enzyme assays. Suggestion of a co-migrating protein was derived from cleavage experiments conducted using PreScission Protease. No sign of degradation was apparent on a Coomassie-stained poly-acrylamide gel despite a large amount of protease (160 U per reaction) and long incubation time (24 h) (data not shown). Removal of GroEL may be achieved by spiking lysates with ATP and GroES;[385] this might be an appropriate experiment in future work.

5.3.10. Experiments with pMH903:PfsufS[M]χ.

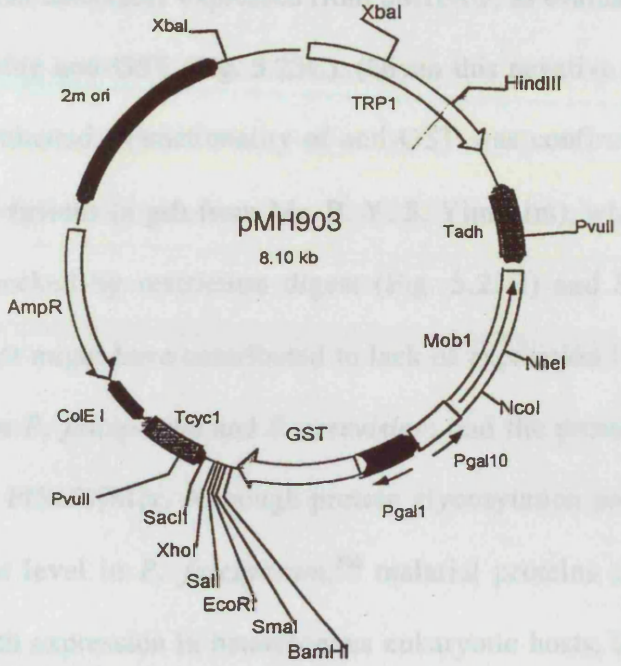
PfSufS[M]χ was cloned into two novel vectors (pMG and pMH903; Fig. 5.23A) developed by Dr. S. G. Sedgwick for heterologous protein expression in *S. cerevisiae*. Protein expressed from both vectors was fused at the N-terminal to GST.

FIGURE 5.23.

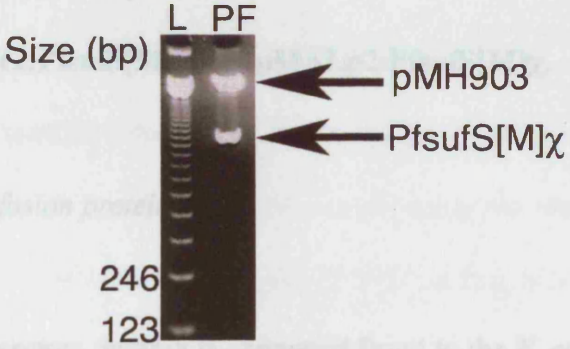
GST-PFSUFS[M]χ EXPRESSION IN *SACCHAROMYCES CEREVISIAE.*

PfSufS[M]χ was cloned into a yeast expression vector (pMH903; A). Insert was confirmed by restriction digest (B) and Sanger sequencing. Total protein from induced yeast cells harbouring the vector was extracted and resolved through NuPage 12 % BT gel. Western blotting using anti-GST indicated that GST-PfSufS[M]χ was not expressed (C). Function of the Western blotting system was confirmed by cross-reactivity against GST-containing irrelevant controls. Abbreviations are as follows. L: 123 bp ladder (Sigma); M: Mark12 unstained standards (Invitrogen); PF: pMH903:PfsufS[M]χ; I1 and I2: irrelevant GST-containing controls.

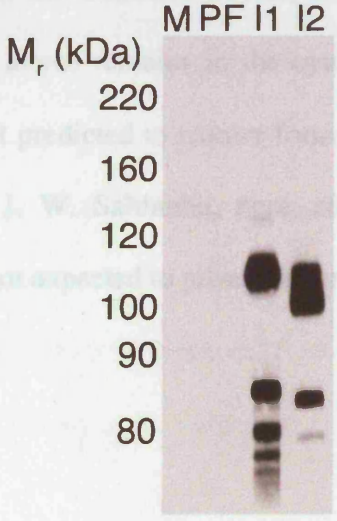
A


XbaI
XbaI
2m ori
TRP1
HindIII
pMH903
8.10 kb
Tadh
PvuII
AmpR
Mob1
NheI
NcoI
ColE I
Tcyc1
GST
Pgal10
PvuII
Pgal1
SacII
XhoI
SalI
EcoRI
SmaI
BamHI

B


L
PF
Size (bp)
pMH903
PfsufS[M]χ
246
123

C


M PF I1 I2
Mr (kDa)
220
160
120
100
90
80

The two vectors were distinct only in the presence of a $3C^{pro}$ site between GST and the insert in pMH903. No protein was detectably expressed from pMH903, as evaluated by Western blot of total protein using anti-GST (Fig. 5.23C). (Given this negative result, pMG:PfsufS[M]χ was not evaluated.) Functionality of anti-GST was confirmed by blotting against irrelevant GST fusions (a gift from Mr. B. Y. S. Yim Lim), while the integrity of the insert was checked by restriction digest (Fig. 5.23B) and Sanger sequencing. Obvious factors that might have contributed to lack of expression include divergent codon usage between *P. falciparum* and *S. cerevisiae*; and the presence of potential glycosylation sites in PfSufS[M]χ. Although protein glycosylation probably does not occur at a significant level in *P. falciparum*,[386] malarial proteins contain potential sites that interfere with expression in heterologous eukaryotic hosts. Indeed, six potential N-linked glycosylation sites were detected in PfSufS[M]χ.

5.3.11. Experiments with pMALc2:PfsufS[M]χ and pMALp2:PfsufS[M]χ.

5.3.11.1. Properties of the fusion protein.

From the pMALc2 and pMALp2 vectors, protein is expressed fused to the *E. coli* MBP encoded by the *malE* gene.[387] In the cMALc2 vector, MBP lacks the periplasmic targeting sequence (pMBP) and hence remains in the cytosol (cMBP). PfSufS[M]χ contains eight cysteines but is not predicted to require formation of disulphide bridges to give a functional fold (Dr. J. W. Saldanha, pers. comm.); thus the reducing environment of the cytoplasm is not expected to present a barrier to folding.

5.3.11.2. Protein expression.

Expression of protein was induced by addition of IPTG to cells growing at 37 °C for 5 h (Fig. 5.24). Cross-reactivity with anti-MBP antibody was marked (Fig. 5.25). Furthermore, soluble protein was present after induction for 18-20 h at 14 °C; protein accumulation was greatest in the BL21-CodonPlus (DE3)-RIL host cells (Fig. 5.25). As determined by image analysis of Coomassie-stained poly-acrylamide gels, the M_r of the cross-reactive protein was 90 kDa; this approximates the predicted size (92 kDa). Therefore, I concluded that soluble cMBP-PfSufS[M]χ was expressed. Conversely, pMBP-PfSufS[M]χ did not appear to be expressed (data not shown); thus, I decided to discontinue experiments with this construct.

5.3.11.3. Affinity purification in a batch format.

Protein cMBP-PfSufS[M]χ was purified over an amylose resin in a batch format by washing and centrifugation (Fig. 5.26). The fusion protein eluted with low specificity. Indeed, poor binding to amylose is associated with the use of MBP as fusion partner.[388] There are also numerous truncation or degradation products.

5.3.11.4. Cleavage experiments.

Efforts were made to cleave cMBP from PfSufS[M]χ by binding protein to amylose resin, prior to incubation with Factor Xa for a prolonged period (78 h). However, evidence for cleavage to yield a PfSufS[M]χ product of *ca.* 50 kDa was unconvincing as viewed on a Coomassie-stained poly-acrylamide gel (Fig. 5.27).

FIGURE 5.24.

EXPRESSION OF cMBP-PFSUFS[M]χ

E. coli cells harbouring pMALc2:PfsufS[M]χ were induced with 0.3 mM IPTG at 37 °C. Samples were taken at 3, 4, and 5 h post-induction. Additionally, an uninduced culture was grown and a sample taken after 5 h of growth. Protein was resolved through NuPage 4-12 % BT gel. Evidence for expression of fusion protein is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); 3, 4 and 5: samples taken 3, 4 and 5 h post-induction; UI: uninduced sample taken at 5 h.

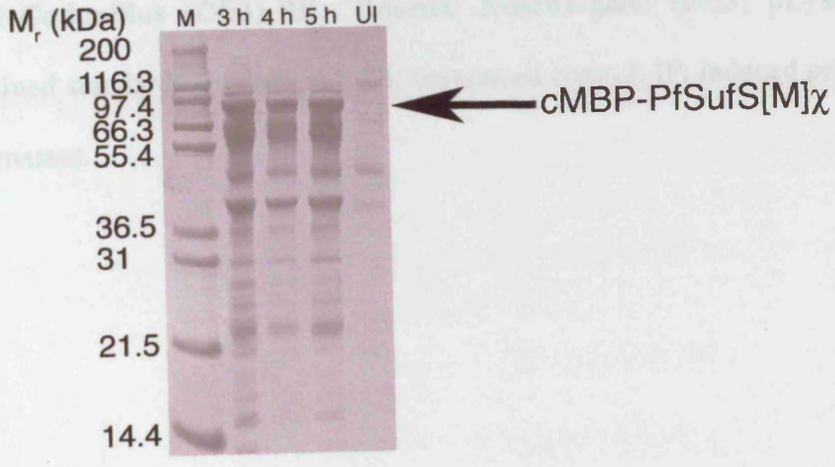

M$_r$ (kDa)
M
3 h
4 h
5 h
UI
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ

FIGURE 5.25.

EXPRESSION OF cMBP-PFSUFS[M]χ AT LOW TEMPERATURE.

Rosetta-gami (DE3) pLysS (Novagen) or BL21-CodonPlus (DE3)-RIL (Stratagene) *E. coli* cells harbouring pMALc2:PfsufS[M]χ were induced with 0.3 mM IPTG at 37 °C for 3 h or 14 °C for 16 h. Protein was extracted in "BugBuster" reagent (Novagen). Soluble protein (supernatant) was fractionated from insoluble protein (pellet) by centrifugation (16000*g*) and resolved through NuPage 12 % BT gel (panel 1). Western blotting was performed using anti-MBP (panel 2). Evidence for expression of soluble cMBP-PfSufS[M]χ is marked with an arrow. Abbreviations are as follows. BL21: BL21-CodonPlus (DE3)-RIL; Rosetta: Rosetta-gami (DE3) pLysS; M: Mark12 unstained standards (Invitrogen); UI: uninduced control; IP: induced pellet; IS: induced supernatant.

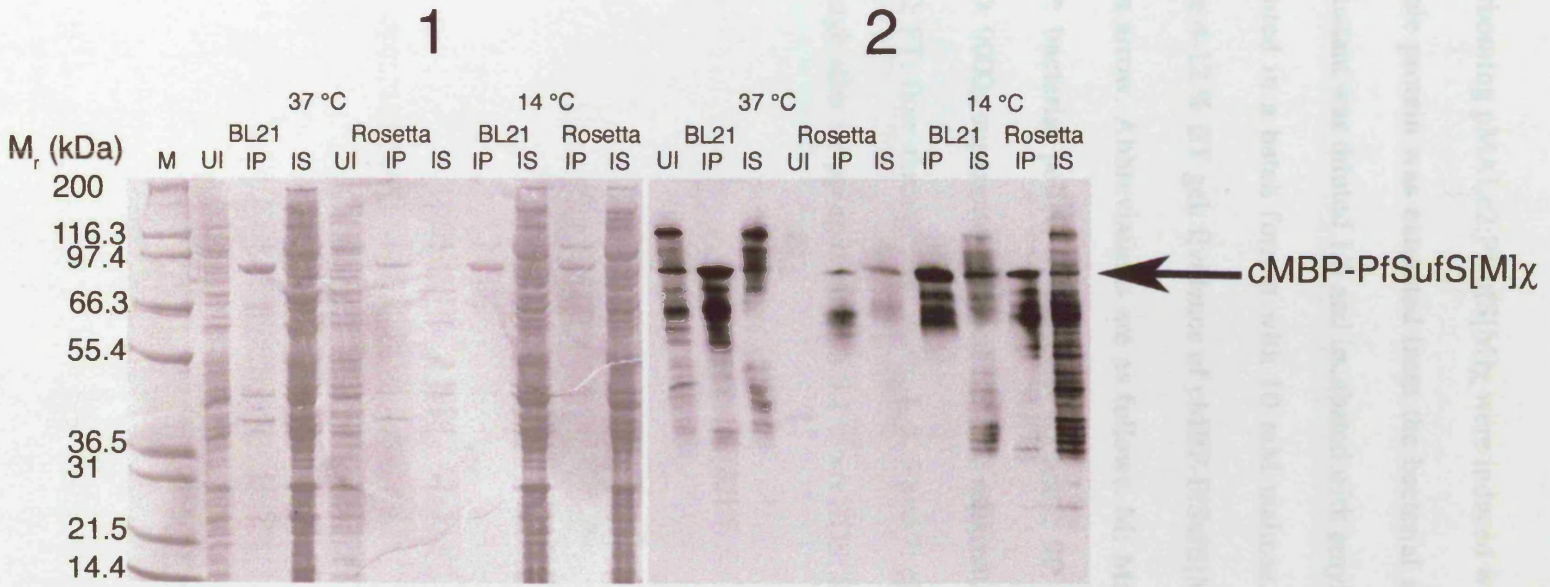

1
2
37 °C
14 °C
37 °C
14 °C
BL21
Rosetta
BL21
Rosetta
BL21
Rosetta
BL21
Rosetta
M
UI
IP
IS
UI
IP
IS
IP
IS
IP
IS
UI
IP
IS
UI
IP
IS
IP
IS
IP
IS
M$_r$ (kDa)
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ

FIGURE 5.26.

PURIFICATION OF cMBP-PFSUFS[M]χ.

E. coli cells harbouring pMALc2:PfsufS[M]χ were induced with 0.3 mM IPTG at 14 °C for 16 h. Soluble protein was extracted from the bacterial sonicate by centrifugation (9000*g*). Supernatant was diluted 1/5 and incubated with amylose resin for 18 h at 4 °C. Protein was eluted in a batch format with 10 mM maltose. Fractions were resolved through NuPage 4-12 % BT gel. Evidence of cMBP-PfSufS[M]χ in elution fraction 1 is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); P: bacterial pellet; S: bacterial sonicate; 9P: 9000*g* pellet; 9S: 9000*g* supernatant; 9D: 9000*g* supernatant diluted 1/5; On: supernatant and resin after binding for 18 h at 4 °C; FT: flow-through; W: wash; 1, 3, 5 and 7: elution fractions 1, 3, 5 and 7; St: flow-through after stripping resin with 0.1 % v/v SDS; SR: stripped resin.

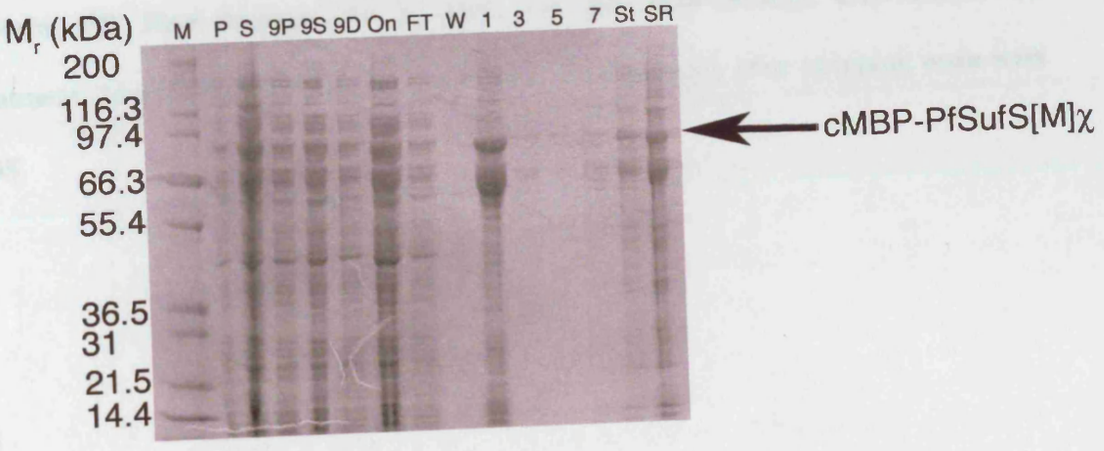
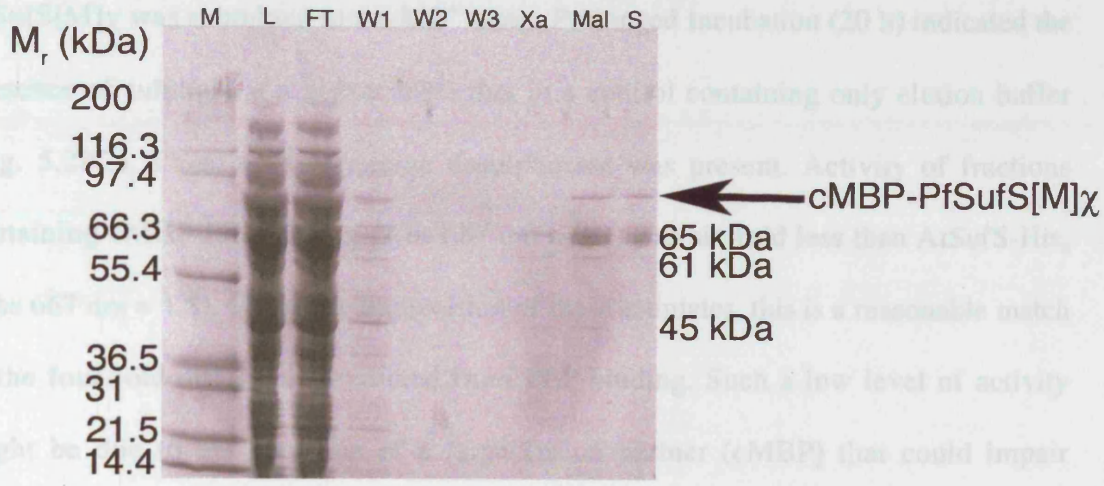

M_r (kDa)
M P S 9P 9S 9D On FT W 1 3 5 7 St SR
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ

FIGURE 5.27.

FACTOR Xa CLEAVAGE OF cMBP-PFSUFS[M]χ.

E. coli cells harbouring pMALc2:PfsufS[M]χ were induced with 0.3 mM IPTG at 14 °C for 16 h. Soluble protein was bound to amylose resin for 2 h at 4 °C. Effort was made to elute the protein by incubation with Factor Xa for 78 h at 4 °C. Remaining protein was eluted with 10 mM maltose followed by 0.1 % v/v SDS. Fractions were resolved through NuPage 4-12 % BT gel. Evidence for cleavage of cMBP-PfSufS[M]χ was unconvincing given the estimated M_r value of cleaved PfSufS[M]χ (50 kDa). Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); L: soluble protein; FT: flow-through; W1-3: wash 1-3; Xa: flow-through after Factor Xa treatment; Mal: flow-through after maltose; S: flow-through after stripping resin with SDS.


M
L
FT
W1
W2
W3
Xa
Mal
S
M$_r$ (kDa)
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ
65 kDa
61 kDa
45 kDa

5.3.11.5. Enzymology.

Elution fractions containing cMBP-PfSufS[M]χ were pooled and concentrated five-fold using a spin concentrator (Fig. 5.28A). Scanning of concentrated protein gave a low (0.26) peak at 408 nm; possibly indicative of Schiff base (Fig. 5.28B). By comparison, an equivalent fraction of AtSufS-His_6 gave a significantly higher peak (0.91) at 419 nm (see chapter 4). All preparations were blanked against redistilled water, ensuring that absorbance data are directly comparable. Therefore, if the amount of Schiff base linkage is used as a crude method of active site titration, it would seem reasonable to predict a four-fold lower activity of cMBP-PfSufS[M]χ than AtSufS-His_6. Concentrated cMBP-PfSufS[M]χ was submitted to the MB^+ assay. Prolonged Incubation (20 h) indicated the presence of sulphide at a higher level that in a control containing only elution buffer (Fig. 5.28C). Thus, active cysteine desulphurase was present. Activity of fractions containing cMBP-PfSufS[M]χ (Abs 667 nm ≈ 0.3) was six-fold less than AtSufS-His_6 (Abs 667 nm ≈ 1.8). Given the imprecision of these estimates, this is a reasonable match to the four-fold difference predicted from PLP binding. Such a low level of activity might be due to the presence of a large fusion partner (cMBP) that could impair dimerization. Reciprocal inhibition might be expected, given the interference by PfSufS[M]χ with the function of cMBP (section 5.3.11.6). Furthermore, optimal conditions for activity of PfSufS[M]χ are unknown.

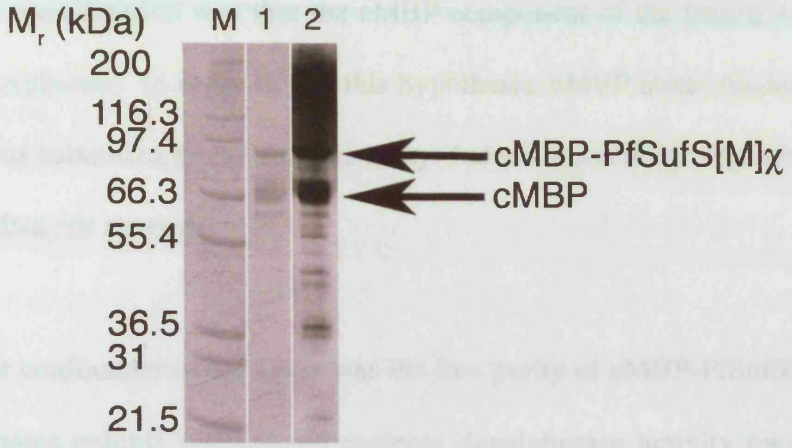
Before proceeding further, it was necessary to investigate numerous potentially confounding factors. The first possibility was that the MB^+ signal was due to sulphides present in the isolate; or to abiotic desulphurase activity of free PLP over prolonged incubation.

FIGURE 5.28.

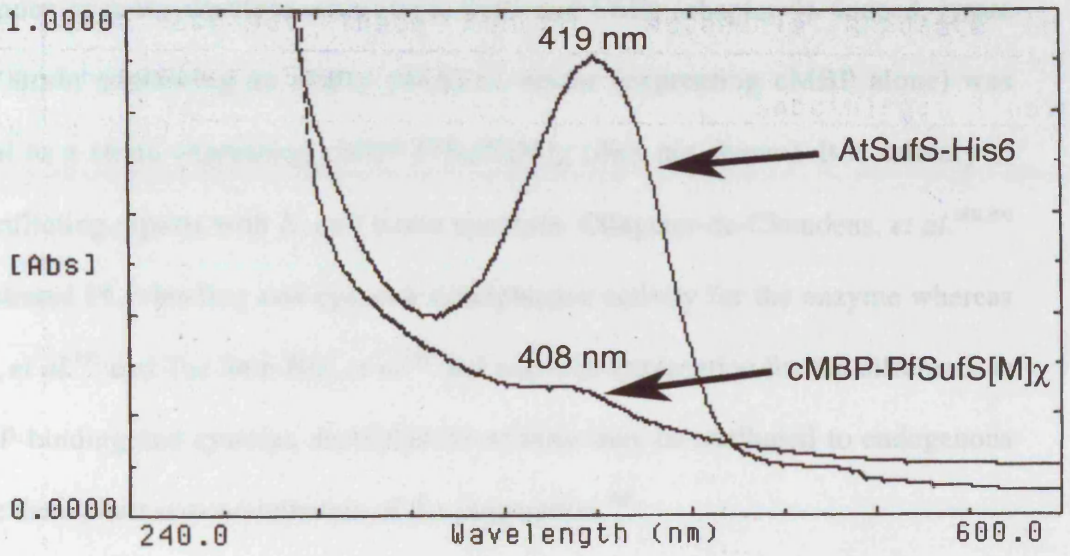
ENZYMATIC CHARACTERISATION OF cMBP-PFSUFS[M]χ.

cMBP-PfSufS[M]χ (Fig. 5.26) was concentrated 5-fold and resolved through NuPage 4-12 % BT gel (A, panel 1) before Western blotting was done with anti-MBP (A, panel 2). A spectral scan was acquired, suggestive of PLP Schiff base (B). A scan of AtSufS-His_6 (chapter 5) is presented for comparison. Cysteine desulphurase activity of the preparation was demonstrated by the MB^+ assay, as compared to a control lacking protein (C).

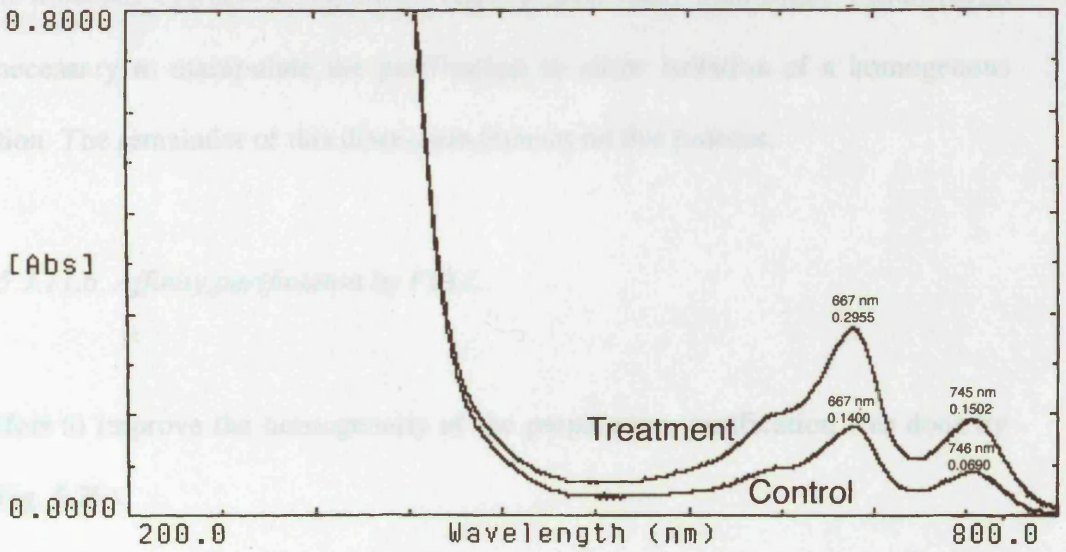
A


M_r (kDa)
M 1 2
200
116.3
97.4
66.3
55.4
36.5
31
21.5
cMBP-PfSufS[M]χ
cMBP

B


1.0000
[Abs]
0.0000
419 nm
AtSufS-His6
408 nm
cMBP-PfSufS[M]χ
240.0
Wavelength (nm)
600.0

C


0.8000
[Abs]
0.0000
667 nm
0.2955
667 nm
0.1400
745 nm
0.1502
746 nm
0.0690
Treatment
Control
200.0
Wavelength (nm)
800.0

However, this possibility can probably be discounted because heat inactivation of the preparation ablated the activity (absorbance spectrum could not be acquired due to high turbidity). Another consideration was that the cMBP component of the fusion was an active cysteine desulphurase. In order to test this hypothesis, cMBP alone (isolated in section 5.3.11.6) was submitted to the enzyme assay. Indeed, no activity was detected with cMBP alone (data not shown).

However, the major confounder in the assay was the low purity of cMBP-PfSufS[M]χ. Crude bacterial lysates exhibit pronounced cysteine desulphurase activity owing to endogenous cysteine desulphurases (IscS, SufS and CSD) (chapter 4). Indeed, lysate from a strain containing an empty pMALc2 vector (expressing cMBP alone) was identical to a strain expressing cMBP-PfSufS[M]χ (data not shown). It is salutary to note conflicting reports with *E. coli* biotin synthase. Ollagnier-de-Choudens, *et al.*[389,390] demonstrated PLP-binding and cysteine desulphurase activity for the enzyme whereas Cosper, *et al.*[391] and Tse Sum Bui, *et al.*[392] did not. One explanation for the difference is that PLP-binding and cysteine desulphurase activity may be attributed to endogenous cysteine desulphurase contamination of the preparation.[393]

Hence, to associate cysteine desulphurase activity specifically with cMBP-PfSufS[M]χ, it was necessary to manipulate the purification to allow isolation of a homogenous preparation. The remainder of this discussion focuses on this process.

5.3.11.6. Affinity purification by FPLC.

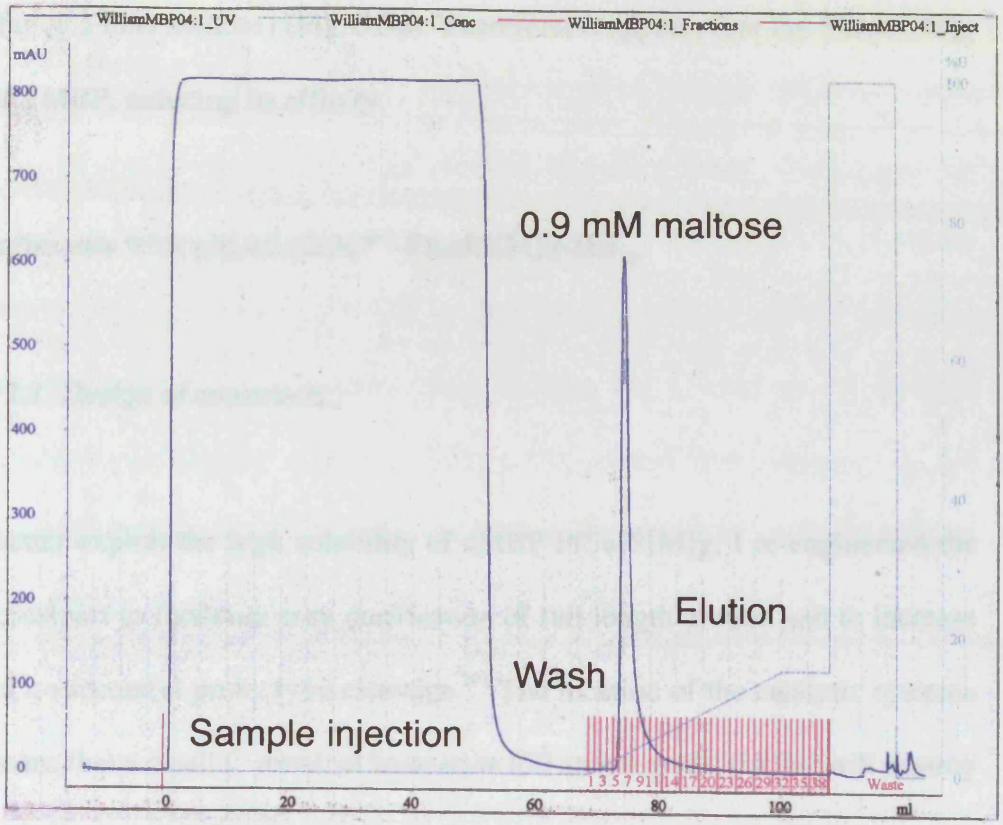
In an effort to improve the homogeneity of the preparation, purification was done by FPLC (Fig. 5.29).

FIGURE 5.29.

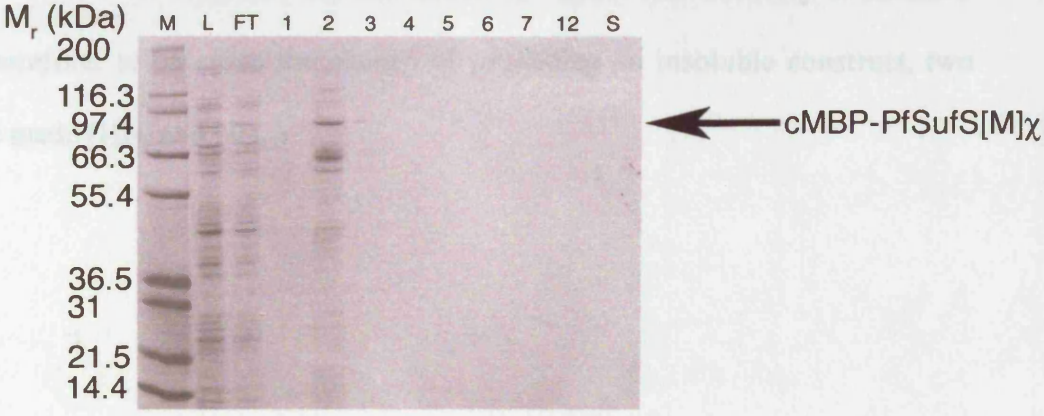
FPLC PURIFICATION OF cMBP-PFSUFS[M]χ

E. coli cells harbouring pMALc2:PfsufS[M]χ were induced with 0.3 mM IPTG at 14 °C for 16 h. Soluble protein was extracted from bacterial sonicate, diluted 1/5, and passed through an amylose resin column by FPLC. Protein was eluted with a linear gradient limited by 10 mM maltose. A typical 280 nm elution profile is presented (A). Fractions were resolved through NuPage 4-12 % BT gel (B). Evidence of the fusion protein in elution fraction 2 is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); L: lysed bacteria; FT: column flow-through; 1-7 and 12: elution fractions 1-7 and 12; S: flow-through after stripping column with 0.1 % v/v SDS.

A


WilliamMBP04:1_UV
WilliamMBP04:1_Conc
WilliamMBP04:1_Fractions
WilliamMBP04:1_Inject
mAU
0.9 mM maltose
Elution
Wash
Sample injection
Waste
ml

B


Mr (kDa)
M L FT 1 2 3 4 5 6 7 12 S
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ

However, the fusion protein again appeared to bind the column weakly and eluted in the dead volume of the apparatus regardless of the gradient used (data not shown). By contrast, cMBP alone bound the column strongly and eluted at a high concentration of limiting buffer (8.3 mM maltose) (Fig. 5.30). Therefore, it appears that the PfSufS[M]χ interacts with cMBP, reducing its affinity.

5.3.12. Experiments with pMALc2:3C^{pro}-PfsufS[M]χ-His$_{12}$.

5.3.12.1. Design of constructs.

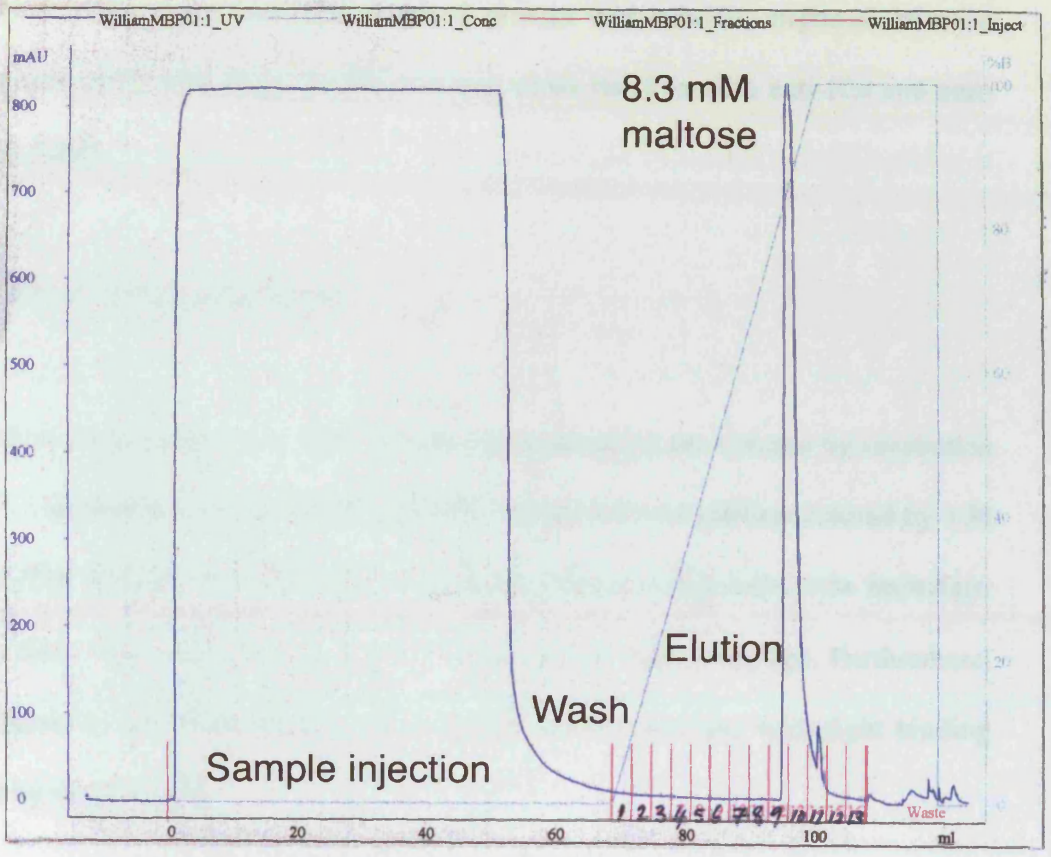
In order to better exploit the high solubility of cMBP-PfSufS[M]χ, I re-engineered the expression construct to facilitate easy purification of full length protein and to increase the chance of a successful proteolytic cleavage.[394] The location of the catalytic cysteine of PfSufS means that a small C-terminal truncation (50 amino acids; 6 kDa) will destroy the activity of the enzyme. Hence, in a bespoke design, a C-terminal affinity tag was introduced to select full-length products during purification. Furthermore, a 3C^{pro} recognition site was inserted between cMBP and PfSufS[M]χ, given the specificity of 3C^{pro} even at low temperature. PCR was used to introduce these modifications. It has been shown that a C-terminal His-tag can markedly reduce the solubility of an MBP fusion.[395] Therefore, to decrease the chance of producing an insoluble construct, two fusions were made (His$_6$ and His$_{12}$)

FIGURE 5.30.

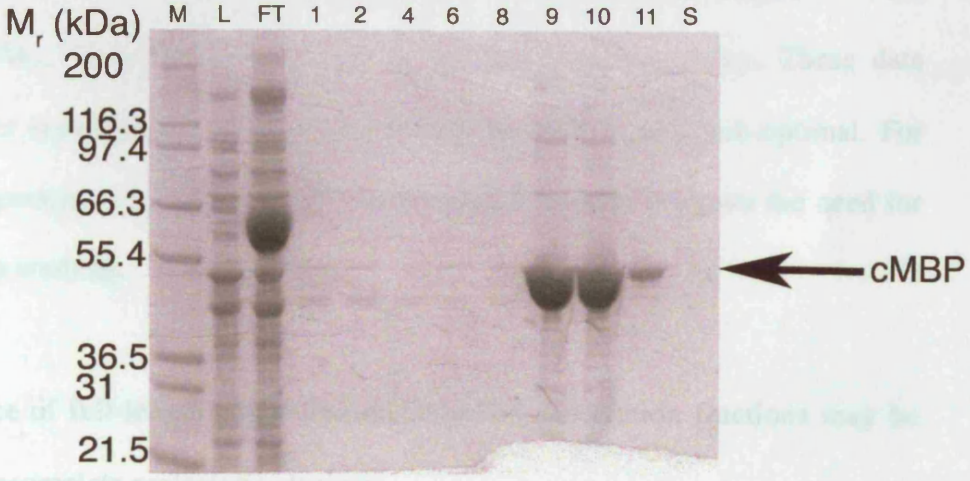
FPLC PURIFICATION OF cMBP.

E. coli cells harbouring empty pMALc2 were induced with 0.3 mM IPTG at 14 °C for 16 h. Soluble protein was extracted from bacterial sonicate, diluted 1/5, and passed through an amylose resin column by FPLC. Protein was eluted with a linear gradient limited by 10 mM maltose. A typical 280 nm elution profile is presented (A). Fractions were resolved through NuPage 4-12 % BT gel (B). Evidence of the protein in elution fractions 9-10 is marked with an arrow. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); L: lysed bacteria; FT: column flow-through; 1-2, 4, 6 and 8-11: elution fractions 1-2, 4, 6 and 8-11; S: flow-through after stripping column with 0.1 % v/v SDS.

A


WilliamMBP01:1_UV
WilliamMBP01:1_Conc
WilliamMBP01:1_Fractions
WilliamMBP01:1_Inject
mAU
800
700
600
500
400
300
200
100
0
8.3 mM
maltose
Elution
Wash
Sample injection
Waste
0
20
40
60
80
100
ml

B


M_r (kDa)
M
L
FT
1
2
4
6
8
9
10
11
S
200
116.3
97.4
66.3
55.4
36.5
31
21.5
cMBP

5.3.12.2. Protein expression.

Modified construct (pMALc2:3C^{pro}-PfSufS[M]χ-His_{12}) was transformed into *E. coli* and expression experiments conducted. Induced protein was soluble, migrated near the predicted size of 95 kDa (Fig. 5.31B) and was cross-reactive with anti-His and anti-MBP (Fig. 5.32).

5.3.12.3. Protein purification.

Soluble extract was applied to a Ni^{2+} column and cleaved on the column by incubation with 3C^{pro}. Cut protein was successfully eluted through a linear gradient limited by 1 M imidazole (Fig. 5.31A). This experiment suggested recovery of protein to be imperfect. However, there was evidence for column binding and protease cleavage. Furthermore, high imidazole (1 M) was required for elution; a result consistent with tight binding conferred by the His_{12}-tag.

Western blotting with anti-His and anti-MBP antibodies (Fig. 5.32) revealed successful cleavage of the fusion protein: products ran near the predicted sizes (cMBP: 44 kDa; PfSufS[M]χ-His_{12}: 51 kDa) and gave the appropriate cross-reactivity. These data warrant further investigation, as digestion conditions were clearly sub-optimal. For example, the presence of cleaved cMBP in the elution fraction suggests the need for more extensive washing.

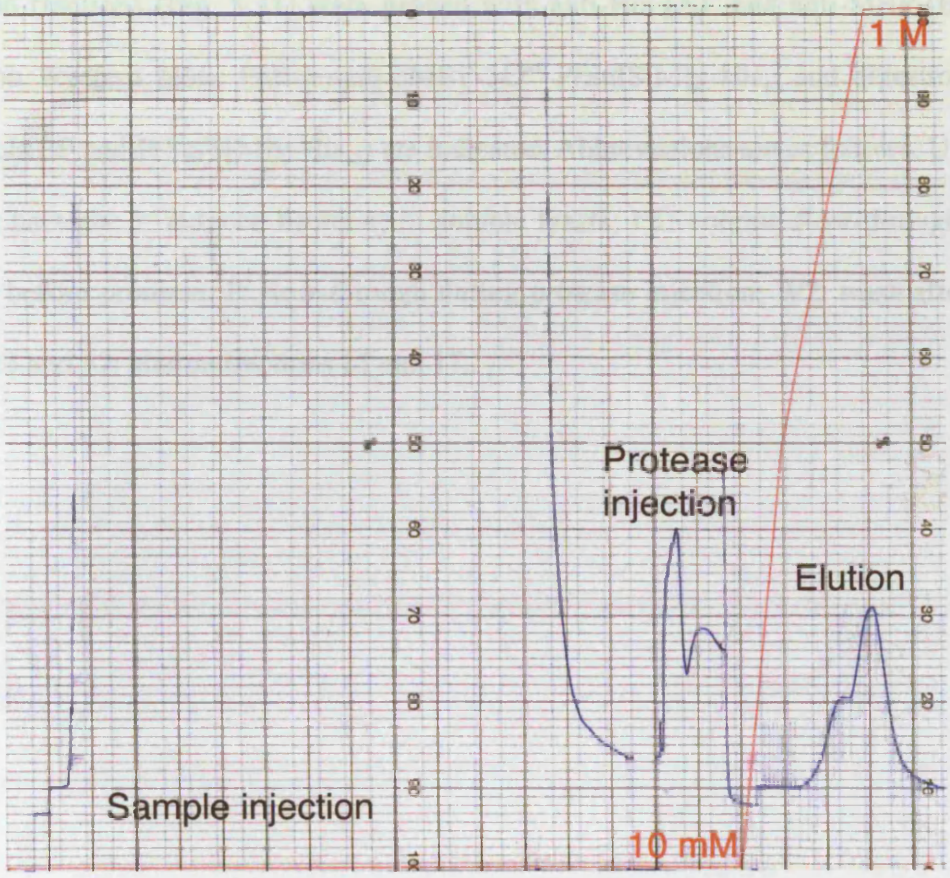
The persistence of full-length cMBP-PfSufS[M]χ-His_{12} in elution fractions may be explained by incomplete proteolytic cleavage.

FIGURE 5.31.

FPLC PURIFICATION OF cMBP-3C^{pro}-PFSUFS[M]χ-HIS$_{12}$.

E. coli cells harbouring pMALc2:3C^{pro}-PfsufS[M]χ-His$_{12}$ were induced with 0.3 mM IPTG at 14 °C for 16 h. Soluble protein was extracted from bacterial sonicate and passed over Ni^{2+} by FPLC. 3C^{pro} (160 U) was injected and the column was incubated at 4 °C for 20 h. Protein was eluted with a linear gradient limited by 1 M imidazole. A typical 280 nm elution profile is presented (A). Fractions were resolved through NuPage 4-12 % BT gel (B). Evidence of the fusion protein in elution fractions 16-18 is presented. Abbreviations are as follows. M: Mark12 unstained standards (Invitrogen); UI: uninduced bacteria; I: induced bacteria; S: soluble lysate; FT: column flow-through; W1: wash before protease; P: flow-through during protease injection; W2: wash after protease; 11-14 and 16-18: elution fractions 11-14 and 16-18.

A


1 M
Protease
injection
Elution
Sample injection
10 mM

B

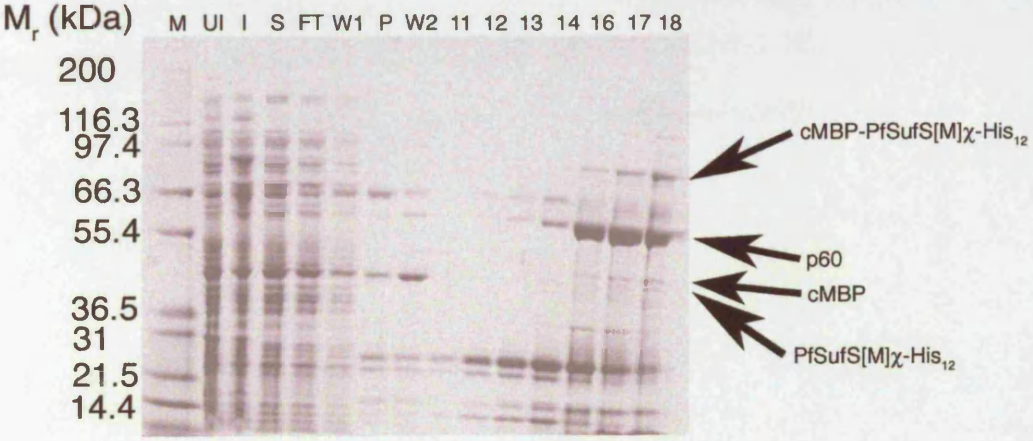
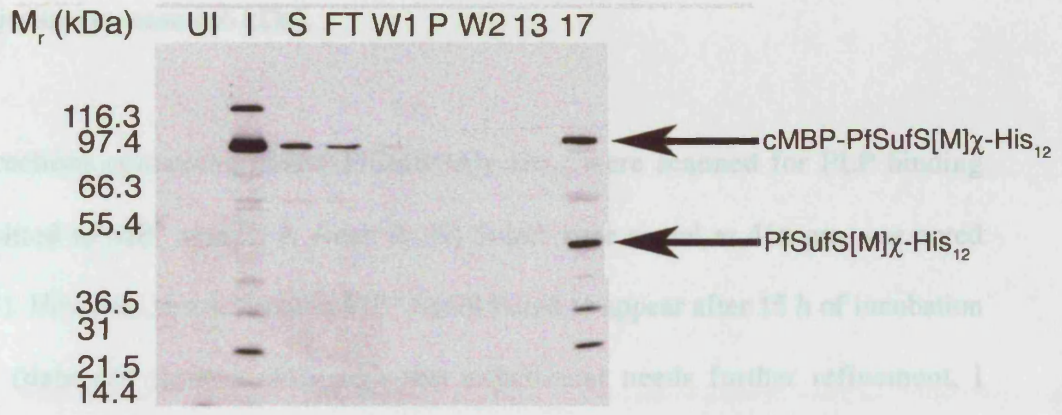

M_r (kDa)
M UI I S FT W1 P W2 11 12 13 14 16 17 18
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ-His12
p60
cMBP
PfSufS[M]χ-His12

FIGURE 5.32.

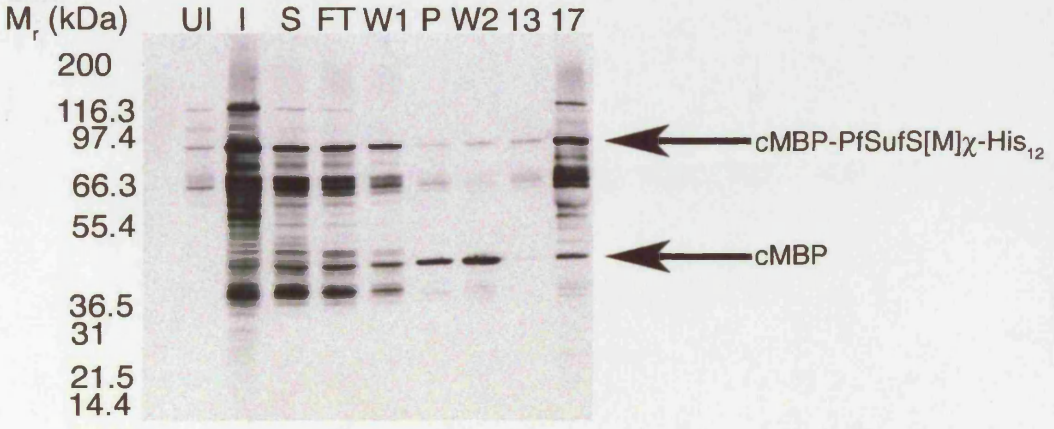
IMMUNO-REACTIVITY OF cMBP-3C^{pro}-PFSUFS[M]χ-HIS$_{12}$.

Purification fractions (Fig. 5.31) were probed with anti-His (A) and anti-MBP (B) antibody on Western blots. Full-length cMBP-3C^{pro}-PfSufS[M]χ-His$_{12}$ and digestion products (cMBP and PfSufS[M]χ-His$_{12}$) are indicated. Abbreviations are as follows. UI: uninduced bacteria; I: induced bacteria; S: soluble lysate; FT: column flow-through; W1: wash before protease; P: flow-through during protease injection; W2: wash after protease; 13 and 17: elution fractions 13 and 17.

A


M$_r$ (kDa)
UI I S FT W1 P W2 13 17
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ-His$_{12}$
PfSufS[M]χ-His$_{12}$

B


M$_r$ (kDa)
UI I S FT W1 P W2 13 17
200
116.3
97.4
66.3
55.4
36.5
31
21.5
14.4
cMBP-PfSufS[M]χ-His$_{12}$
cMBP

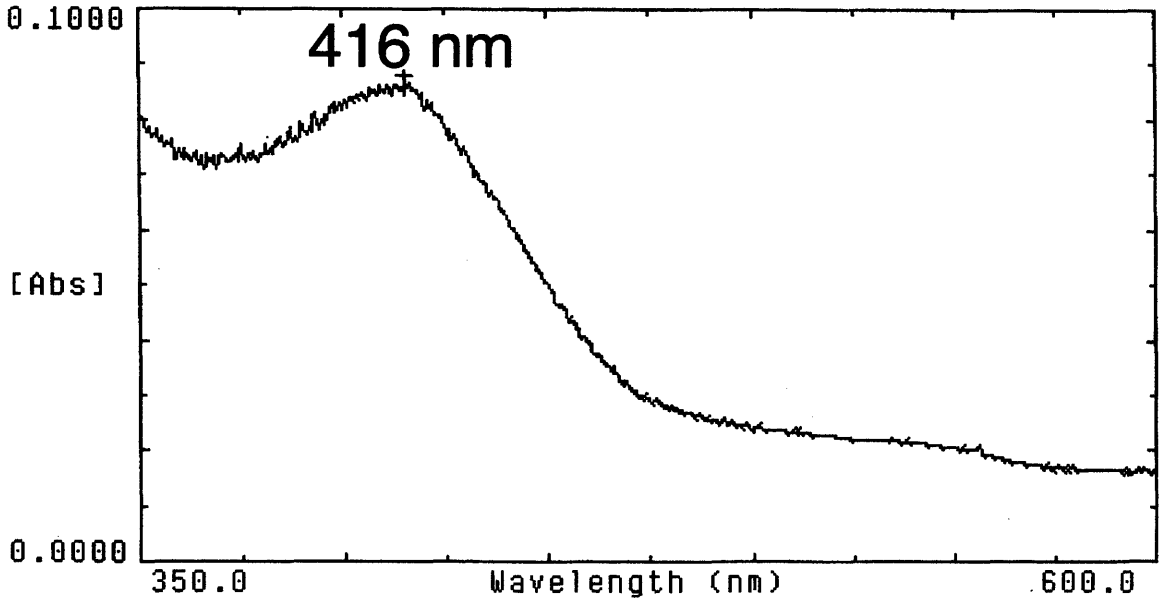
This phenomenon is probably due to low effective protease concentration in the column, given the inability of $3C^{pro}$ to adhere to the Ni^{2+} resin. A more efficient fusion protease would be 6His-$3C^{pro}$ (Ref. [381]), as it could be immobilised in the resin. It is also worth noting the co-elution of a significant amount of an unidentified protein of *ca.* 60 kDa and lacking cross-reactivity to anti-His_{12} and anti-MBP antibodies. On the basis of size, amount and elution profile (high imidazole), it is reasonable to exclude this being the PreScission protease (46 kDa).

Elution fractions containing cMBP-PfSufS[M]χ-His_{12} were scanned for PLP binding and submitted to MB^+ assays. A weak (0.09) Schiff base signal at 416 nm was noted (Fig. 5.33). However, a colorimetric MB^+ signal failed to appear after 15 h of incubation at 37 °C (data not shown). Although this experiment needs further refinement, I tentatively suggest that the cMBP-PfSufS[M]χ-His_{12} is enzymatically inactive and that the initial activity described (section 5.3.12.5) was due to contaminants. However, cleavage of PfSufS[M]χ (away from cMBP) might render the enzyme able to dimerize and become active.

FIGURE 5.33.

ABSORBANCE SPECTRUM OF cMBP-3C^{pro}-PFSUFS[M]χ-HIS$_{12}$.

An absorbance spectrum of elution fraction 17 (Fig. 5.31) was acquired. A weak (0.09) Schiff base peak is apparent at 416 nm.


0.1000
416 nm
[Abs]
0.0000
350.0
Wavelength (nm)
600.0

5.3.13. Future approach.

Expression of soluble *P. falciparum* protein in a heterologous host cell is known to be difficult, e.g., recent work conducted in this laboratory failed to produce active *P. falciparum* SufA, SufB, SufC and PBGS (Rangachari, Clough, Thakrar and Wilson, NIMR, unpublished observations). However, by consideration of the nature of the PfSufS sequence; coupled to simple experiments with solubility tags, it has been possible to obtain soluble protein (cMBP-PfSufS[M]χ). Furthermore, by bespoke re-engineering of this construct, a soluble fusion amenable to cleavage and facile purification (cMBP-PfSufS[M]χ-His_{12}) was produced. A promising route is now open to evaluate the enzymatic activity of PfSufS.

GENERAL CONCLUSION.

The objective of this project was to investigate a sufS gene in *P. falciparum.* In order to achieve this, a number of difficult techniques have been applied and progress towards this objective has been made. Using a combination of simple bioinformatic and *in vitro* methods, the existence of the gene and its transcription has been demonstrated in *P. falciparum.* A start has been made on investigation of PfSufS localisation using transfection and immunocytochemistry. The major focus of this project was the production of active recombinant PfSufS to allow a full description of enzymatic properties. A robust colorimetric cysteine desulphurase assay based on methylene blue has been developed and significant progress has been made towards the production of recombinant enzyme. This work may have a practical purpose as large amounts of enzyme and a rapid assay would be required for high throughput *in vitro* screening of inhibitors. Future efforts will capitalise on these achievements, as we move towards a complete characterisation of the intracellular localisation and enzymatic properties of PfSufS.

REFERENCES.

1. Hackett, L. W. *Malaria in Europe: an ecological study* (Humphrey Milford, London, 1937).
2. Snow, R. W., Guerra, C. A., Noor, A. M., Myint, H. Y. & Hay, S. I. The global distribution of clinical episodes of Plasmodium falciparum malaria. *Nature* **434**, 214-217 (2005).
3. Sherman, I. W. in *Malaria: parasite biology, pathogenesis, and protection* (ed. Sherman, I. W.) (American Society for Microbiology, Washington, DC, 1998).
4. Ascenzi, A. Malaria diseases and parasites. *Parassitologia* **41**, 33-8 (1999).
5. Nye, E. R. Alphonse Laveran (1845-1922): discoverer of the malarial parasite and Nobel laureate, 1907. *J Med Biogr* **10**, 81-7 (2002).
6. Cavalier-Smith, T. Protist phylogeny and the high-level classification of Protozoa. *Europ J Protistol* **39**, 338-348 (2003).
7. Perkins, F. O., Barta, J. R., Clopton, R. E., Peirce, M. A. & Upton, S. J. in *An Illustrated Guide to the Protozoa* (eds. Lee, J. J., Leedale, G. F. & Bradbury, P.) (Society of Protozoologists, Lawrence, KA, 2000).
8. Escalante, A. A. & Ayala, F. J. Phylogeny of the malarial genus Plasmodium, derived from rRNA gene sequences. *Proc Natl Acad Sci U S A* **91**, 11373-7 (1994).
9. Su, X. Z., Mu, J. & Joy, D. A. The "Malaria's Eve" hypothesis and the debate concerning the origin of the human malaria parasite Plasmodium falciparum. *Microbes Infect* **5**, 891-6 (2003).
10. Rich, S. M., Licht, M. C., Hudson, R. R. & Ayala, F. J. Malaria's Eve: evidence of a recent population bottleneck throughout the world populations of Plasmodium falciparum. *Proc Natl Acad Sci U S A* **95**, 4425-30 (1998).
11. Mu, J. et al. Chromosome-wide SNPs reveal an ancient origin for Plasmodium falciparum. *Nature* **418**, 323-6 (2002).
12. Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. & Mitchell, G. H. A brief illustrated guide to the ultrastructure of Plasmodium falciparum asexual blood stages. *Parasitol Today* **16**, 427-33 (2000).
13. Markell, E. K., John, D. T. & Krotoski, W. A. *Markell and Voge's Medical Parasitology* (W. B. Saunders Company, Philadelphia, 1999).
14. Honigsbaum, M. *The Fever Trail: the Hunt for the Cure for Malaria* (Macmillan, London, 2001).
15. Frevert, U. Sneaking in through the back entrance: the biology of malaria liver stages. *Trends Parasitol* **20**, 417-24 (2004).
16. Bannister, L. & Mitchell, G. The ins, outs and roundabouts of malaria. *Trends Parasitol* **19**, 209-13 (2003).
17. Trager, W. & Jensen, J. B. Continuous culture of Plasmodium falciparum: its impact on malaria research. *Int J Parasitol* **27**, 989-1006 (1997).
18. Bozdech, Z. et al. The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum. *PLoS Biol* **1**, E5 (2003).
19. Beier, J. C. et al. Quantitation of malaria sporozoites transmitted in vitro during salivation by wild Afrotropical Anopheles. *Med Vet Entomol* **5**, 71-9 (1991).
20. Shortt, H. E., Fairley, N. H., Covell, G., Shute, P. G. & Garnham, P. C. The pre-erythrocytic stage of Plasmodium falciparum. *Trans R Soc Trop Med Hyg* **44**, 405-19 (1951).
21. Talman, A. M., Domarle, O., McKenzie, F. E., Ariey, F. & Robert, V. Gametocytogenesis: the puberty of Plasmodium falciparum. *Malar J* **3**, 24 (2004).

22. Miller, L. H., Baruch, D. I., Marsh, K. & Doumbo, O. K. The pathogenic basis of malaria. *Nature* **415**, 673-9 (2002).
23. Warrell, D. A. Management of severe malaria. *Parassitologia* **41**, 287-94 (1999).
24. Bosman, A., Delacollette, C., Olumese, P., Ridley, R. G. & Rietveld, A. (WHO, Geneva, 2001).
25. Li, Y. & Wu, Y. L. An over four millennium story behind qinghaosu (artemisinin)--a fantastic antimalarial drug from a traditional chinese herb. *Curr Med Chem* **10**, 2197-230 (2003).
26. Haynes, R. K. & Krishna, S. Artemisinins: activities and actions. *Microbes Infect* **6**, 1339-46 (2004).
27. Fitch, C. D. Ferriprotoporphyrin IX, phospholipids, and the antimalarial actions of quinoline drugs. *Life Sci* **74**, 1957-72 (2004).
28. Dobson, M. J. The malariology centenary. *Parassitologia* **41**, 21-32 (1999).
29. Bynum, W. F. Portraits of science. Mosquitoes bite more than once. *Science* **295**, 47-8 (2002).
30. Fantini, B. The concept of specificity and the Italian contribution to the discovery of the malaria transmission cycle. *Parassitologia* **41**, 39-47 (1999).
31. Coluzzi, M. Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control. *Bull World Health Organ* **62 Suppl**, 107-13 (1984).
32. Marquardt, W. C., Demaree, R. S. & Grieve, R. B. *Parasitology and Vector Biology* (Academic Press, San Diego, CA, 2000).
33. Sinden, R. E. Plasmodium differentiation in the mosquito. *Parassitologia* **41**, 139-48 (1999).
34. Osta, M. A., Christophides, G. K., Vlachou, D. & Kafatos, F. C. Innate immunity in the malaria vector Anopheles gambiae: comparative and functional genomics. *J Exp Biol* **207**, 2551-63 (2004).
35. Najera, J. A. Malaria control: achievements, problems and strategies. *Parassitologia* **43**, 1-89 (2001).
36. Greenwood, B. What can the residents of malaria endemic countries do to protect themselves against malaria? *Parassitologia* **41**, 295-9 (1999).
37. Bruce-Chwatt, L. J. & de Zulueta, J. *The Rise and Fall of Malaria in Europe* (Oxford University Press, Oxford, 1980).
38. White, N. J. Antimalarial drug resistance. *J Clin Invest* **113**, 1084-92 (2004).
39. Hemingway, J. & Ranson, H. Insecticide resistance in insect vectors of human disease. *Annu Rev Entomol* **45**, 371-91 (2000).
40. Ridley, R. G. Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* **415**, 686-93 (2002).
41. Richie, T. L. & Saul, A. Progress and challenges for malaria vaccines. *Nature* **415**, 694-701 (2002).
42. Alphey, L. et al. Malaria control with genetically manipulated insect vectors. *Science* **298**, 119-21 (2002).
43. Varmus, H. et al. Public health. Enhanced: Grand Challenges in Global Health. *Science* **302**, 398-9 (2003).
44. Venter, J. C. et al. The sequence of the human genome. *Science* **291**, 1304-51 (2001).
45. Gardner, M. J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. *Nature* **419**, 498-511 (2002).
46. Holt, R. A. et al. The genome sequence of the malaria mosquito Anopheles gambiae. *Science* **298**, 129-49 (2002).

47. Ballou, W. R. et al. Update on the clinical development of candidate malaria vaccines. *Am J Trop Med Hyg* **71**, 239-47 (2004).
48. Jacobs-Lorena, M. Interrupting malaria transmission by genetic manipulation of anopheline mosquitoes. *J Vector Borne Dis* **40**, 73-7 (2003).
49. Biagini, G. A., O'Neill, P. M., Nzila, A., Ward, S. A. & Bray, P. G. Antimalarial chemotherapy: young guns or back to the future? *Trends Parasitol* **19**, 479-87 (2003).
50. Scholtyseck, E. & Piekarski, G. [Electron microscopic studies on merozoites of Eimeria (Eimeria perforans and E. stidae) and Toxoplasma gondii. On the systematic position of T. gondii]. *Z Parasitenkd* **26**, 91-115 (1965).
51. Gutteridge, W. E., Trigg, P. I. & Williamson, D. H. Properties of DNA from some malarial parasites. *Parasitology* **62**, 209-19 (1971).
52. Wilson, R. J., Williamson, D. H. & Preiser, P. Malaria and other Apicomplexans: the "plant" connection. *Infect Agents Dis* **3**, 29-37 (1994).
53. Wilson, R. J. et al. Complete gene map of the plastid-like DNA of the malaria parasite Plasmodium falciparum. *J Mol Biol* **261**, 155-72 (1996).
54. Kohler, S. et al. A plastid of probable green algal origin in Apicomplexan parasites. *Science* **275**, 1485-9 (1997).
55. McFadden, G. I., Reith, M. E., Munholland, J. & Lang-Unnasch, N. Plastid in human parasites. *Nature* **381**, 482 (1996).
56. Cai, X., Fuller, A. L., McDougald, L. R. & Zhu, G. Apicoplast genome of the coccidian Eimeria tenella. *Gene* **321**, 39-46 (2003).
57. Zhu, G., Marchewka, M. J. & Keithly, J. S. Cryptosporidium parvum appears to lack a plastid genome. *Microbiology* **146 (Pt 2)**, 315-21 (2000).
58. Sudman, M. S. Protothecosis. A critical review. *Am J Clin Pathol* **61**, 10-9 (1974).
59. Hopkins, J. et al. The plastid in Plasmodium falciparum asexual blood stages: a three-dimensional ultrastructural analysis. *Protist* **150**, 283-95 (1999).
60. Waller, R. F. & McFadden, G. I. The apicoplast: a review of the derived plastid of apicomplexan parasites. *Curr Issues Mol Biol* **7**, 57-79 (2005).
61. Zhang, Z., Green, B. R. & Cavalier-Smith, T. Phylogeny of ultra-rapidly evolving dinoflagellate chloroplast genes: a possible common origin for sporozoan and dinoflagellate plastids. *J Mol Evol* **51**, 26-40 (2000).
62. Fast, N. M., Kissinger, J. C., Roos, D. S. & Keeling, P. J. Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids. *Mol Biol Evol* **18**, 418-26 (2001).
63. Dyall, S. D., Brown, M. T. & Johnson, P. J. Ancient invasions: from endosymbionts to organelles. *Science* **304**, 253-7 (2004).
64. Delwiche, C. F. Tracing the Thread of Plastid Diversity through the Tapestry of Life. *Am Nat* **154**, S164-S177 (1999).
65. Yoon, H. S., Hackett, J. D., Pinto, G. & Bhattacharya, D. The single, ancient origin of chromist plastids. *Proc Natl Acad Sci U S A* **99**, 15507-12 (2002).
66. Funes, S., Reyes-Prieto, A., Perez-Martinez, X. & Gonzalez-Halphen, D. On the evolutionary origins of apicoplasts: revisiting the rhodophyte vs. chlorophyte controversy. *Microbes Infect* **6**, 305-11 (2004).
67. Blanchard, J. L. & Hicks, J. S. The non-photosynthetic plastid in malarial parasites and other apicomplexans is derived from outside the green plastid lineage. *J Eukaryot Microbiol* **46**, 367-75 (1999).
68. Coppin, A. et al. Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite Toxoplasma gondii argues for a red alga ancestry. *J Mol Evol* **60**, 257-67 (2005).
69. Funes, S. et al. A green algal apicoplast ancestor. *Science* **298**, 2155 (2002).

70. Waller, R. F., Keeling, P. J., van Dooren, G. G. & McFadden, G. I. Comment on "A green algal apicoplast ancestor". *Science* **301**, 49; author reply 49 (2003).
71. Obornik, M. et al. Phylogenetic analyses suggest lateral gene transfer from the mitochondrion to the apicoplast. *Gene* **285**, 109-18 (2002).
72. Zhu, X. Y. Phylogenetic analysis indicates bacteria-to-apicoplast lateral gene transfer. *Yi Chuan Xue Bao* **31**, 1316-20 (2004).
73. Huang, C. Y., Ayliffe, M. A. & Timmis, J. N. Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature* **422**, 72-6 (2003).
74. Stegemann, S., Hartmann, S., Ruf, S. & Bock, R. High-frequency gene transfer from the chloroplast genome to the nucleus. *Proc Natl Acad Sci U S A* **100**, 8828-33 (2003).
75. Lister, D. L., Bateman, J. M., Purton, S. & Howe, C. J. DNA transfer from chloroplast to nucleus is much rarer in Chlamydomonas than in tobacco. *Gene* **316**, 33-8 (2003).
76. van Dooren, G. G., Waller, R. F., Joiner, K. A., Roos, D. S. & McFadden, G. I. Traffic jams: protein transport in Plasmodium falciparum. *Parasitol Today* **16**, 421-7 (2000).
77. van Dooren, G. G., Su, V., D'Ombrain, M. C. & McFadden, G. I. Processing of an apicoplast leader sequence in Plasmodium falciparum and the identification of a putative leader cleavage enzyme. *J Biol Chem* **277**, 23612-9 (2002).
78. Zuegge, J., Ralph, S., Schmuker, M., McFadden, G. I. & Schneider, G. Deciphering apicoplast targeting signals--feature extraction from nuclear-encoded precursors of Plasmodium falciparum apicoplast proteins. *Gene* **280**, 19-26 (2001).
79. Foth, B. J. et al. Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. *Science* **299**, 705-8 (2003).
80. Ralph, S. A. et al. Tropical infectious diseases: Metabolic maps and functions of the Plasmodium falciparum apicoplast. *Nat Rev Microbiol* **2**, 203-16 (2004).
81. Striepen, B. et al. The plastid of Toxoplasma gondii is divided by association with the centrosomes. *J Cell Biol* **151**, 1423-34 (2000).
82. Sinden, R. E. & Strong, K. An ultrastructural study of the sporogonic development of Plasmodium falciparum in Anopheles gambiae. *Trans R Soc Trop Med Hyg* **72**, 477-91 (1978).
83. Le Roch, K. G. et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science* **301**, 1503-8 (2003).
84. Daily, J. P. et al. In Vivo Transcriptome of Plasmodium falciparum Reveals Overexpression of Transcripts That Encode Surface Proteins. *J Infect Dis* **191**, 1196-203 (2005).
85. Le Roch, K. G. et al. Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle. *Genome Res* **14**, 2308-18 (2004).
86. Florens, L. et al. A proteomic view of the Plasmodium falciparum life cycle. *Nature* **419**, 520-6 (2002).
87. Lasonder, E. et al. Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. *Nature* **419**, 537-42 (2002).
88. Wang, Q., Brown, S., Roos, D. S., Nussenzweig, V. & Bhanot, P. Transcriptome of axenic liver stages of Plasmodium yoelii. *Mol Biochem Parasitol* **137**, 161-8 (2004).
89. Hall, N. et al. A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. *Science* **307**, 82-6 (2005).
90. Neuhaus, H. E. & Emes, M. J. Nonphotosynthetic Metabolism in Plastids. *Annu Rev Plant Physiol Plant Mol Biol* **51**, 111-140 (2000).

91. Williamson, D. H. et al. The plastid DNA of the malaria parasite Plasmodium falciparum is replicated by two mechanisms. *Mol Microbiol* **45**, 533-42 (2002).
92. Roy, A., Cox, R. A., Williamson, D. H. & Wilson, R. J. Protein synthesis in the plastid of Plasmodium falciparum. *Protist* **150**, 183-8 (1999).
93. Chaubey, S., Kumar, A., Singh, D. & Habib, S. The apicoplast of Plasmodium falciparum is translationally active. *Mol Microbiol* **56**, 81-9 (2005).
94. Fischer, K. & Weber, A. Transport of carbon in non-green plastids. *Trends Plant Sci* **7**, 345-51 (2002).
95. Vollmer, M., Thomsen, N., Wiek, S. & Seeber, F. Apicomplexan parasites possess distinct nuclear-encoded, but apicoplast-localized, plant-type ferredoxin-NADP+ reductase and ferredoxin. *J Biol Chem* **276**, 5483-90 (2001).
96. Surolia, N. & Surolia, A. Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of Plasmodium falciparum. *Nat Med* **7**, 167-73 (2001).
97. Waller, R. F. et al. Nuclear-encoded proteins target to the plastid in Toxoplasma gondii and Plasmodium falciparum. *Proc Natl Acad Sci U S A* **95**, 12352-7 (1998).
98. Jelenska, J. et al. Subcellular localization of acetyl-CoA carboxylase in the apicomplexan parasite Toxoplasma gondii. *Proc Natl Acad Sci U S A* **98**, 2723-8 (2001).
99. Foth, B. J. et al. The malaria parasite Plasmodium falciparum has only one pyruvate dehydrogenase complex, which is located in the apicoplast. *Mol Microbiol* **55**, 39-53 (2005).
100. Wrenger, C. & Muller, S. The human malaria parasite Plasmodium falciparum has distinct organelle-specific lipoylation pathways. *Mol Microbiol* **53**, 103-13 (2004).
101. Thomsen-Zieger, N., Schachtner, J. & Seeber, F. Apicomplexan parasites contain a single lipoic acid synthase located in the plastid. *FEBS Lett* **547**, 80-6 (2003).
102. Gronwald, J. W. Herbicides inhibiting acetyl-CoA carboxylase. *Biochem Soc Trans* **22**, 616-21 (1994).
103. Waller, R. F. et al. A type II pathway for fatty acid biosynthesis presents drug targets in Plasmodium falciparum. *Antimicrob Agents Chemother* **47**, 297-301 (2003).
104. Cassera, M. B. et al. The Methylerythritol Phosphate Pathway Is Functionally Active in All Intraerythrocytic Stages of Plasmodium falciparum. *J Biol Chem* **279**, 51749-59 (2004).
105. Eisenreich, W., Rohdich, F. & Bacher, A. Deoxyxylulose phosphate pathway to terpenoids. *Trends Plant Sci* **6**, 78-84 (2001).
106. Roos, D. S. et al. Mining the Plasmodium genome database to define organellar function: what does the apicoplast do? *Philos Trans R Soc Lond B Biol Sci* **357**, 35-46 (2002).
107. Jomaa, H. et al. Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**, 1573-6 (1999).
108. Tonkin, C. J. et al. Localization of organellar proteins in Plasmodium falciparum using a novel set of transfection vectors and a new immunofluorescence fixation method. *Mol Biochem Parasitol* **137**, 13-21 (2004).
109. Sato, S., Clough, B., Coates, L. & Wilson, R. J. Enzymes for heme biosynthesis are found in both the mitochondrion and plastid of the malaria parasite Plasmodium falciparum. *Protist* **155**, 117-25 (2004).

110. Varadharajan, S., Sagar, B. K., Rangarajan, P. N. & Padmanaban, G. Localization of ferrochelatase in Plasmodium falciparum. *Biochem J* **384**, 429-36 (2004).
111. Sato, S. & Wilson, R. J. The genome of Plasmodium falciparum encodes an active delta-aminolevulinic acid dehydratase. *Curr Genet* **40**, 391-8 (2002).
112. Sato, S. & Wilson, R. J. Proteobacteria-like ferrochelatase in the malaria parasite. *Curr Genet* **42**, 292-300 (2003).
113. Fichera, M. E. & Roos, D. S. A plastid organelle as a drug target in apicomplexan parasites. *Nature* **390**, 407-9 (1997).
114. Clough, B. & Wilson, R. J. M. in *Antimalarial Chemotherapy: mechanisms of action, resistance, and new directions in drug discovery* (ed. Rosenthal, P. J.) (Humana Press, Totowa, NJ, 2001).
115. Camps, M., Arrizabalaga, G. & Boothroyd, J. An rRNA mutation identifies the apicoplast as the target for clindamycin in Toxoplasma gondii. *Mol Microbiol* **43**, 1309-18 (2002).
116. WHO. (World Health Organisation, Geneva, 2005).
117. Gornicki, P. Apicoplast fatty acid biosynthesis as a target for medical intervention in apicomplexan parasites. *Int J Parasitol* **33**, 885-96 (2003).
118. Model, E. & Bindler, J. (Ciba-Geigy AG, Switzerland, 1966).
119. Perozzo, R. et al. Structural elucidation of the specificity of the antibacterial agent triclosan for malarial enoyl acyl carrier protein reductase. *J Biol Chem* **277**, 13106-14 (2002).
120. McLeod, R. et al. Triclosan inhibits the growth of Plasmodium falciparum and Toxoplasma gondii by inhibition of apicomplexan Fab I. *Int J Parasitol* **31**, 109-13 (2001).
121. Samuel, B. U. et al. Delivery of antimicrobials into parasites. *Proc Natl Acad Sci U S A* **100**, 14281-6 (2003).
122. Rodriguez-Concepcion, M. The MEP pathway: a new target for the development of herbicides, antibiotics and antimalarial drugs. *Curr Pharm Des* **10**, 2391-400 (2004).
123. Okuhara, M. et al. Studies on new phosphonic acid antibiotics. III. Isolation and characterization of FR-31564, FR-32863 and FR-33289. *J Antibiot (Tokyo)* **33**, 24-8 (1980).
124. Missinou, M. A. et al. Fosmidomycin for malaria. *Lancet* **360**, 1941-2 (2002).
125. Lell, B. et al. Fosmidomycin, a novel chemotherapeutic agent for malaria. *Antimicrob Agents Chemother* **47**, 735-8 (2003).
126. Borrmann, S. et al. Fosmidomycin-clindamycin for Plasmodium falciparum Infections in African children. *J Infect Dis* **189**, 901-8 (2004).
127. Borrmann, S. et al. Fosmidomycin-clindamycin for the treatment of Plasmodium falciparum malaria. *J Infect Dis* **190**, 1534-40 (2004).
128. Beinert, H., Holm, R. H. & Munck, E. Iron-sulfur clusters: nature's modular, multipurpose structures. *Science* **277**, 653-9 (1997).
129. Seemann, M. et al. Isoprenoid biosynthesis through the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase (GcpE) is a [4Fe-4S] protein. *Angew Chem Int Ed Engl* **41**, 4337-9 (2002).
130. Wolff, M. et al. Isoprenoid biosynthesis via the methylerythritol phosphate pathway: the (E)-4-hydroxy-3-methylbut-2-enyl diphosphate reductase (LytB/IspH) from Escherichia coli is a [4Fe-4S] protein. *FEBS Lett* **541**, 115-20 (2003).
131. Seeber, F. Biogenesis of iron-sulphur clusters in amitochondriate and apicomplexan protists. *Int J Parasitol* **32**, 1207-17 (2002).

132. Crane, F. L. Role of trace elements in electron transport and oxidative phosphorylation. *Soil Sci* **85**, 78-86 (1958).
133. Beinert, H. & Sands, R. H. Studies on succinic and DPNH dehydrogenase preparations by paramagnetic resonance (EPR) spectroscopy. *Biochem Biophys Res Commun* **3**, 41-46 (1960).
134. Sands, R. H. & Beinert, H. Studies on mitochondria and submitochondrial particles by paramagnetic resonance (EPR) spectroscopy. *Biochem Biophys Res Commun* **3**, 47-52 (1960).
135. Beinert, H. Spectroscopy of succinate dehydrogenases, a historical perspective. *Biochim Biophys Acta* **1553**, 7-22 (2002).
136. Frausto da Silva, J. J. R. & Williams, R. J. P. *The Biological Chemistry of the Elements: the Inorganic Chemistry of Life* (Clarendon Press, Oxford, 1991).
137. Jacobson, M. R. et al. Biochemical and genetic analysis of the nifUSVWZM cluster from Azotobacter vinelandii. *Mol Gen Genet* **219**, 49-57 (1989).
138. Muhlenhoff, U. & Lill, R. Biogenesis of iron-sulfur proteins in eukaryotes: a novel task of mitochondria that is inherited from bacteria. *Biochim Biophys Acta* **1459**, 370-82 (2000).
139. Frazzon, J., Fick, J. R. & Dean, D. R. Biosynthesis of iron-sulphur clusters is a complex and highly conserved process. *Biochem Soc Trans* **30**, 680-5 (2002).
140. Gerber, J. & Lill, R. Biogenesis of iron-sulfur proteins in eukaryotes: components, mechanism and pathology. *Mitochondrion* **2**, 71-86 (2002).
141. Balk, J. & Lill, R. The cell's cookbook for iron--sulfur clusters: recipes for fool's gold? *Chembiochem* **5**, 1044-9 (2004).
142. Mansy, S. S. & Cowan, J. A. Iron-sulfur cluster biosynthesis: toward an understanding of cellular machinery and molecular mechanism. *Acc Chem Res* **37**, 719-25 (2004).
143. Venkateswara Rao, P. & Holm, R. H. Synthetic analogues of the active sites of iron-sulfur proteins. *Chem Rev* **104**, 527-59 (2004).
144. Natarajan, K. & Cowan, J. A. Identification of a key intermediate of relevance to iron-sulphur cluster biosynthesis. Mechanism of cluster assembly and implications for protein folding. *J Am Chem Soc* **119**, 4082-4083 (1997).
145. Kennedy, M. C., Emptage, M. H., Dreyer, J. L. & Beinert, H. The role of iron in the activation-inactivation of aconitase. *J Biol Chem* **258**, 11098-105 (1983).
146. Djaman, O., Outten, F. W. & Imlay, J. A. Repair of oxidized iron-sulfur clusters in Escherichia coli. *J Biol Chem* **279**, 44590-9 (2004).
147. Keyer, K. & Imlay, J. A. Inactivation of dehydratase [4Fe-4S] clusters and disruption of iron homeostasis upon cell exposure to peroxynitrite. *J Biol Chem* **272**, 27652-9 (1997).
148. Rogers, P. A. & Ding, H. L-cysteine-mediated destabilization of dinitrosyl iron complexes in proteins. *J Biol Chem* **276**, 30980-6 (2001).
149. Tsai, M. L. et al. Photochemistry of the dinitrosyl iron complex [S5Fe(NO)2]- leading to reversible formation of [S5Fe(mu-S)2FeS5]2-: spectroscopic characterization of species relevant to the nitric oxide modification and repair of [2Fe-2S] ferredoxins. *Inorg Chem* **43**, 5159-67 (2004).
150. Yang, W., Rogers, P. A. & Ding, H. Repair of nitric oxide-modified ferredoxin [2Fe-2S] cluster by cysteine desulfurase (IscS). *J Biol Chem* **277**, 12868-73 (2002).
151. Ollagnier-de-Choudens, S., Sanakis, Y. & Fontecave, M. SufA/IscA: reactivity studies of a class of scaffold proteins involved in [Fe-S] cluster assembly. *J Biol Inorg Chem* **9**, 828-38 (2004).

152. Tokumoto, U. et al. Network of protein-protein interactions among iron-sulfur cluster assembly proteins in Escherichia coli. *J Biochem (Tokyo)* **131**, 713-9 (2002).
153. Muhlenhoff, U., Richhardt, N., Gerber, J. & Lill, R. Characterization of iron-sulfur protein assembly in isolated mitochondria. A requirement for ATP, NADH, and reduced iron. *J Biol Chem* **277**, 29810-6 (2002).
154. Bilder, P. W., Ding, H. & Newcomer, M. E. Crystal structure of the ancient, Fe-S scaffold IscA reveals a novel protein fold. *Biochemistry* **43**, 133-9 (2004).
155. Cupp-Vickery, J. R., Silberg, J. J., Ta, D. T. & Vickery, L. E. Crystal structure of IscA, an iron-sulfur cluster assembly protein from Escherichia coli. *J Mol Biol* **338**, 127-37 (2004).
156. Bertini, I., Cowan, J. A., Del Bianco, C., Luchinat, C. & Mansy, S. S. Thermotoga maritima IscU. Structural characterization and dynamics of a new class of metallochaperone. *J Mol Biol* **331**, 907-24 (2003).
157. Liochev, S. I. The mechanism of "Fenton-like" reactions and their importance for biological systems. A biologist's view. *Met Ions Biol Syst* **36**, 1-39 (1999).
158. Yoon, T. & Cowan, J. A. Iron-sulfur cluster biosynthesis. Characterization of frataxin as an iron donor for assembly of [2Fe-2S] clusters in ISU-type proteins. *J Am Chem Soc* **125**, 6078-84 (2003).
159. Bulteau, A. L. et al. Frataxin acts as an iron chaperone protein to modulate mitochondrial aconitase activity. *Science* **305**, 242-5 (2004).
160. Mihara, H. & Esaki, N. Bacterial cysteine desulfurases: their function and mechanisms. *Appl Microbiol Biotechnol* **60**, 12-23 (2002).
161. Greenberg, J. T. & Demple, B. Glutathione in Escherichia coli is dispensable for resistance to H2O2 and gamma radiation. *J Bacteriol* **168**, 1026-9 (1986).
162. Zheng, L., White, R. H., Cash, V. L., Jack, R. F. & Dean, D. R. Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. *Proc Natl Acad Sci U S A* **90**, 2754-8 (1993).
163. Zheng, L., Cash, V. L., Flint, D. H. & Dean, D. R. Assembly of iron-sulfur clusters. Identification of an iscSUA-hscBA-fdx gene cluster from Azotobacter vinelandii. *J Biol Chem* **273**, 13264-72 (1998).
164. Mihara, H., Kurihara, T., Yoshimura, T., Soda, K. & Esaki, N. Cysteine sulfinate desulfinase, a NIFS-like protein of Escherichia coli with selenocysteine lyase and cysteine desulfurase activities. Gene cloning, purification, and characterization of a novel pyridoxal enzyme. *J Biol Chem* **272**, 22417-24 (1997).
165. Mihara, H. et al. A nifS-like gene, csdB, encodes an Escherichia coli counterpart of mammalian selenocysteine lyase. Gene cloning, purification, characterization and preliminary x-ray crystallographic studies. *J Biol Chem* **274**, 14768-72 (1999).
166. Patzer, S. I. & Hantke, K. SufS is a NifS-like protein, and SufD is necessary for stability of the [2Fe-2S] FhuF protein in Escherichia coli. *J Bacteriol* **181**, 3307-9 (1999).
167. Kambampati, R. & Lauhon, C. T. IscS is a sulfurtransferase for the in vitro biosynthesis of 4-thiouridine in Escherichia coli tRNA. *Biochemistry* **38**, 16561-8 (1999).
168. Lauhon, C. T. & Kambampati, R. The iscS gene in Escherichia coli is required for the biosynthesis of 4-thiouridine, thiamin, and NAD. *J Biol Chem* **275**, 20096-103 (2000).
169. Skovran, E. & Downs, D. M. Metabolic defects caused by mutations in the isc gene cluster in Salmonella enterica serovar typhimurium: implications for thiamine synthesis. *J Bacteriol* **182**, 3896-903 (2000).

170. Soda, K., Oikawa, T. & Esaki, N. Vitamin B6 enzymes participating in selenium amino acid metabolism. *Biofactors* **10**, 257-62 (1999).
171. Ali, V., Shigeta, Y., Tokumoto, U., Takahashi, Y. & Nozaki, T. An intestinal parasitic protist, Entamoeba histolytica, possesses a non-redundant nitrogen fixation-like system for iron-sulfur cluster assembly under anaerobic conditions. *J Biol Chem* **279**, 16863-74 (2004).
172. Johnson, D. C., Dos Santos, P. C. & Dean, D. R. NifU and NifS are required for the maturation of nitrogenase and cannot replace the function of isc-gene products in Azotobacter vinelandii. *Biochem Soc Trans* **33**, 90-3 (2005).
173. Kurihara, T., Mihara, H., Kato, S., Yoshimura, T. & Esaki, N. Assembly of iron-sulfur clusters mediated by cysteine desulfurases, IscS, CsdB and CSD, from Escherichia coli. *Biochim Biophys Acta* **1647**, 303-9 (2003).
174. Takahashi, Y. & Tokumoto, U. A third bacterial system for the assembly of iron-sulfur clusters with homologs in archaea and plastids. *J Biol Chem* **277**, 28380-3 (2002).
175. Zheng, M. et al. DNA microarray-mediated transcriptional profiling of the Escherichia coli response to hydrogen peroxide. *J Bacteriol* **183**, 4562-70 (2001).
176. Schwartz, C. J. et al. IscR, an Fe-S cluster-containing transcription factor, represses expression of Escherichia coli genes encoding Fe-S cluster assembly proteins. *Proc Natl Acad Sci U S A* **98**, 14895-900 (2001).
177. Zheng, M. & Storz, G. Redox sensing by prokaryotic transcription factors. *Biochem Pharmacol* **59**, 1-6 (2000).
178. Outten, F. W., Djaman, O. & Storz, G. A suf operon requirement for Fe-S cluster assembly during iron starvation in Escherichia coli. *Mol Microbiol* **52**, 861-72 (2004).
179. Nachin, L., El Hassouni, M., Loiseau, L., Expert, D. & Barras, F. SoxR-dependent response to oxidative stress and virulence of Erwinia chrysanthemi: the key role of SufC, an orphan ABC ATPase. *Mol Microbiol* **39**, 960-72 (2001).
180. Wexler, M. et al. Fur is not the global regulator of iron uptake genes in Rhizobium leguminosarum. *Microbiology* **149**, 1357-65 (2003).
181. Todd, J. D., Sawers, G. & Johnston, A. W. Proteomic analysis reveals the wide-ranging effects of the novel, iron-responsive regulator RirA in Rhizobium leguminosarum bv. viciae. *Mol Genet Genomics* (2005).
182. Leon, S., Touraine, B., Ribot, C., Briat, J. F. & Lobreaux, S. Iron-sulphur cluster assembly in plants: distinct NFU proteins in mitochondria and plastids from Arabidopsis thaliana. *Biochem J* **371**, 823-30 (2003).
183. Leon, S., Touraine, B., Briat, J. F. & Lobreaux, S. Mitochondrial localization of Arabidopsis thaliana Isu Fe-S scaffold proteins. *FEBS Lett* **579**, 1930-4 (2005).
184. Devlin, R. M. *Plant Physiology* (D. Van Nostrand Company, New York, 1975).
185. Xu, X. M., Adams, S., Chua, N. H. & Moller, S. G. AtNAP1 represents an atypical SufB protein in Arabidopsis plastids. *J Biol Chem* **280**, 6648-54 (2005).
186. Ollagnier-de Choudens, S. et al. SufA from Erwinia chrysanthemi. Characterization of a scaffold protein required for iron-sulfur cluster assembly. *J Biol Chem* **278**, 17993-8001 (2003).
187. Duraffourg, N. et al. Assignment of the (1)H, (15)N and (13)C resonances of SufA from Escherichia coli involved in Fe-S cluster biosynthesis. *J Biomol NMR* **30**, 379-80 (2004).
188. Law, A. E., Mullineaux, C. W., Hirst, E. M., Saldanha, J. & Wilson, R. J. Bacterial orthologues indicate the malarial plastid gene ycf24 is essential. *Protist* **151**, 317-27 (2000).

189. Nachin, L., Loiseau, L., Expert, D. & Barras, F. SufC: an unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress. *Embo J* **22**, 427-37 (2003).
190. Rangachari, K. et al. SufC hydrolyzes ATP and interacts with SufB from Thermotoga maritima. *FEBS Lett* **514**, 225-8 (2002).
191. Outten, F. W., Wood, M. J., Munoz, F. M. & Storz, G. The SufE protein and the SufBCD complex enhance SufS cysteine desulfurase activity as part of a sulfur transfer pathway for Fe-S cluster assembly in Escherichia coli. *J Biol Chem* **278**, 45713-9 (2003).
192. Loiseau, L., Ollagnier-De-Choudens, S., Nachin, L., Fontecave, M. & Barras, F. Biogenesis of Fe-S cluster by the bacterial Suf system : SufS and SufE form a new type of cysteine desulfurase. *J Biol Chem* (2003).
193. Ollagnier-de-Choudens, S. et al. Mechanistic studies of the SufS-SufE cysteine desulfurase: evidence for sulfur transfer from SufS to SufE. *FEBS Lett* **555**, 263-7 (2003).
194. Goldsmith-Fischman, S. et al. The SufE sulfur-acceptor protein contains a conserved core structure that mediates interdomain interactions in a variety of redox protein complexes. *J Mol Biol* **344**, 549-65 (2004).
195. Touraine, B. et al. Nfu2: a scaffold protein required for [4Fe-4S] and ferredoxin iron-sulphur cluster assembly in Arabidopsis chloroplasts. *Plant J* **40**, 101-11 (2004).
196. Moller, S. G., Kunkel, T. & Chua, N. H. A plastidic ABC protein involved in intercompartmental communication of light signaling. *Genes Dev* **15**, 90-103 (2001).
197. Xu, X. M. & Moller, S. G. AtNAP7 is a plastidic SufC-like ATP-binding cassette/ATPase essential for Arabidopsis embryogenesis. *Proc Natl Acad Sci U S A* **101**, 9143-8 (2004).
198. Leon, S., Touraine, B., Briat, J. F. & Lobreaux, S. The AtNFS2 gene from Arabidopsis thaliana encodes a NifS-like plastidial cysteine desulphurase. *Biochem J* **366**, 557-64 (2002).
199. Pilon-Smits, E. A. et al. Characterization of a NifS-Like Chloroplast Protein from Arabidopsis. Implications for Its Role in Sulfur and Selenium Metabolism. *Plant Physiol* **130**, 1309-18 (2002).
200. Takahashi, Y., Mitsui, A., Fujita, Y. & Matsubara, H. Roles of ATP and NADPH in formation of the Fe-S cluster of spinach ferredoxin. *Plant Physiol* **95**, 104-110 (1991).
201. Takahashi, Y., Mitsui, A. & Matsubara, H. Formation of the Fe-S cluster of ferredoxin in lysed spinach chloroplasts. *Plant Physiol* **95**, 97-103 (1991).
202. Halliwell, B. & Gutteridge, J. M. C. *Free Radicals in Biology and Medicine* (Clarendon Press, Oxford, 1989).
203. Yabe, T. et al. The Arabidopsis chloroplastic NifU-like protein CnfU, which can act as an iron-sulfur cluster scaffold protein, is required for biogenesis of ferredoxin and photosystem I. *Plant Cell* **16**, 993-1007 (2004).
204. Hjorth, E., Hadfi, K., Zauner, S. & Maier, U. G. Unique genetic compartmentalization of the SUF system in cryptophytes and characterization of a SufD mutant in Arabidopsis thaliana. *FEBS Lett* **579**, 1129-35 (2005).
205. Ye, H. et al. The chloroplast NifS-like protein of Arabidopsis thaliana is required for iron-sulfur cluster formation in ferredoxin. *Planta* **220**, 602-8 (2005).
206. Seidler, A., Jaschkowitz, K. & Wollenberg, M. Incorporation of iron-sulphur clusters in membrane-bound proteins. *Biochem Soc Trans* **29**, 418-21 (2001).

207. Carlton, J. M. et al. Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii. *Nature* **419**, 512-9 (2002).
208. Bahl, A. et al. PlasmoDB: the Plasmodium genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished). *Nucleic Acids Res* **30**, 87-90 (2002).
209. Xu, P. et al. The genome of Cryptosporidium hominis. *Nature* **431**, 1107-12 (2004).
210. Abrahamsen, M. S. et al. Complete genome sequence of the apicomplexan, Cryptosporidium parvum. *Science* **304**, 441-5 (2004).
211. Puiu, D., Enomoto, S., Buck, G. A., Abrahamsen, M. S. & Kissinger, J. C. CryptoDB: the Cryptosporidium genome resource. *Nucleic Acids Res* **32**, D329-31 (2004).
212. Kissinger, J. C., Gajria, B., Li, L., Paulsen, I. T. & Roos, D. S. ToxoDB: accessing the Toxoplasma gondii genome. *Nucleic Acids Res* **31**, 234-6 (2003).
213. Conway, D. J. et al. Origin of Plasmodium falciparum malaria is traced by mitochondrial DNA. *Mol Biochem Parasitol* **111**, 163-71 (2000).
214. Rosario, V. Cloning of naturally occurring mixed infections of malaria parasites. *Science* **212**, 1037-8 (1981).
215. Ponnudurai, T., Lensen, A. H., Leeuwenberg, A. D. & Meuwissen, J. H. Cultivation of fertile Plasmodium falciparum gametocytes in semi-automated systems. 1. Static cultures. *Trans R Soc Trop Med Hyg* **76**, 812-8 (1982).
216. Pennisi, E. Bioinformatics. Gene counters struggle to get the right answer. *Science* **301**, 1040-1 (2003).
217. Hayward, R. E. et al. Shotgun DNA microarrays and stage-specific gene expression in Plasmodium falciparum malaria. *Mol Microbiol* **35**, 6-14 (2000).
218. Ben Mamoun, C. et al. Co-ordinated programme of gene expression during asexual intraerythrocytic development of the human malaria parasite Plasmodium falciparum revealed by microarray analysis. *Mol Microbiol* **39**, 26-36 (2001).
219. Bozdech, Z. et al. Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray. *Genome Biol* **4**, R9 (2003).
220. Kappe, S. H. et al. Exploring the transcriptome of the malaria sporozoite stage. *Proc Natl Acad Sci U S A* **98**, 9895-900 (2001).
221. Gunasekera, A. M., Patankar, S., Schug, J., Eisen, G. & Wirth, D. F. Drug-induced alterations in gene expression of the asexual blood forms of Plasmodium falciparum. *Mol Microbiol* **50**, 1229-39 (2003).
222. Isokpehi, R. D. & Hide, W. A. Integrative analysis of intraerythrocytic differentially expressed transcripts yields novel insights into the biology of Plasmodium falciparum. *Malar J* **2**, 38 (2003).
223. Sato, S., Rangachari, K. & Wilson, R. J. Targeting GFP to the malarial mitochondrion. *Mol Biochem Parasitol* **130**, 155-8 (2003).
224. LaGier, M. J., Tachezy, J., Stejskal, F., Kutisova, K. & Keithly, J. S. Mitochondrial-type iron-sulfur cluster biosynthesis genes (IscS and IscU) in the apicomplexan Cryptosporidium parvum. *Microbiology* **149**, 3519-30 (2003).
225. Ellis, K. E., Clough, B., Saldanha, J. W. & Wilson, R. J. Nifs and Sufs in malaria. *Mol Microbiol* **41**, 973-81 (2001).
226. Jansen, R., Greenbaum, D. & Gerstein, M. Relating whole-genome expression data with protein-protein interactions. *Genome Res* **12**, 37-46 (2002).

227. Bozdech, Z. & Ginsburg, H. Data mining of the transcriptome of Plasmodium falciparum: The pentose phosphate pathway and ancillary processes. *Malar J* **4**, 17 (2005).
228. Ng, W. V. et al. Genome sequence of Halobacterium species NRC-1. *Proc Natl Acad Sci U S A* **97**, 12176-81 (2000).
229. Page, R. D. M. & Holmes, E. C. *Molecular evolution: a phylogenetic approach* (Blackwell Science, Oxford, 1998).
230. Hide, W. A. & Isokpehi, R. D. Positive selection scanning of parasite DNA sequences. *Methods Mol Biol* **270**, 127-50 (2004).
231. Pizzi, E. & Frontali, C. Divergence of noncoding sequences and of insertions encoding nonglobular domains at a genomic region well conserved in plasmodia. *J Mol Evol* **50**, 474-80 (2000).
232. Pizzi, E. & Frontali, C. Low-complexity regions in Plasmodium falciparum proteins. *Genome Res* **11**, 218-29 (2001).
233. Kyte, J. & Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32 (1982).
234. Daily, J. P. et al. In vivo transcriptional profiling of Plasmodium falciparum. *Malar J* **3**, 30 (2004).
235. Liochev, S. I. & Fridovich, I. Superoxide and iron: partners in crime. *IUBMB Life* **48**, 157-61 (1999).
236. Wilson, R. J. Parasite plastids: approaching the endgame. *Biol Rev Camb Philos Soc* **80**, 129-53 (2005).
237. Mary, A., Kidd, G. L. & Gross, S. R. The response time of transcription and translation of the leu-2 gene of Neurospora to its inducer, alpha-isopropylmalate, approaches the permissible minimum. *Biochem Biophys Res Commun* **161**, 1286-90 (1989).
238. Scandalios, J. G. Oxidative stress responses--what have genome-scale studies taught us? *Genome Biol* **3**, REVIEWS1019 (2002).
239. Muller, S. Redox and antioxidant systems of the malaria parasite Plasmodium falciparum. *Mol Microbiol* **53**, 1291-305 (2004).
240. Hihara, Y., Sonoike, K., Kanehisa, M. & Ikeuchi, M. DNA microarray analysis of redox-responsive genes in the genome of the cyanobacterium Synechocystis sp. strain PCC 6803. *J Bacteriol* **185**, 1719-25 (2003).
241. Loyevsky, M. et al. Chelation of iron within the erythrocytic Plasmodium falciparum parasite by iron chelators. *Mol Biochem Parasitol* **101**, 43-59 (1999).
242. van Zyl, R. L., Havlik, I. & Monteagudo, F. S. The combined effect of iron chelators and classical antimalarials on the in-vitro growth of Plasmodium falciparum. *J Antimicrob Chemother* **30**, 273-8 (1992).
243. Gordeuk, V. R. & Loyevsky, M. Antimalarial effect of iron chelators. *Adv Exp Med Biol* **509**, 251-72 (2002).
244. Loyevsky, M. et al. An IRP-like protein from Plasmodium falciparum binds to a mammalian iron-responsive element. *Blood* **98**, 2555-62 (2001).
245. Goldman, N. & Yang, Z. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol Biol Evol* **11**, 725-36 (1994).
246. Yang, Z. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* **15**, 555-556 (1997).
247. Bailey, T. L. & Elkan, C. in *Proceedings of the Second International Conference on Intelligent Systems for Molecular Biology* (eds. Altman, R., Brutlag, D., Karp, P., Lathrop, R. & Searls, D.) 28-36 (AAAI Press, Menlo Park, CA, 1994).
248. Notredame, C., Higgins, D. G. & Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**, 205-17 (2000).

249. Chenna, R. et al. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res* **31**, 3497-500 (2003).
250. Jones, D. T., Taylor, W. R. & Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* **8**, 275-82 (1992).
251. Lima, C. D. Analysis of the E. coli NifS CsdB protein at 2.0 A reveals the structural basis for perselenide and persulfide intermediate formation. *J Mol Biol* **315**, 1199-208 (2002).
252. Jensen, J. B. in *Malaria Methods and Protocols* (ed. Doolan, D. L.) (Humana Press, Totowa, NJ, 2002).
253. Hempelmann, E., Ling, I. & Wilson, R. J. S-antigens and isozymes in strains of Plasmodium falciparum. *Trans R Soc Trop Med Hyg* **75**, 855-8 (1981).
254. Voss, T. in *Malaria Methods and Protocols* (ed. Doolan, D. L.) (Humana Press, Totowa, NJ, 2002).
255. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Press, Cold Spring Harbor, 1989).
256. Wesseling, J. G., Dirks, R., Smits, M. A. & Schoenmakers, J. G. Nucleotide sequence and expression of a beta-tubulin gene from Plasmodium falciparum, a malarial parasite of man. *Gene* **83**, 301-9 (1989).
257. van Dooren, G. G., Schwartzbach, S. D., Osafune, T. & McFadden, G. I. Translocation of proteins across the multiple membranes of complex plastids. *Biochim Biophys Acta* **1541**, 34-53 (2001).
258. Diniz, J. A., Silva, E. O., Lainson, R. & de Souza, W. The fine structure of Garnia gonadati and its association with the host cell. *Parasitol Res* **86**, 971-7 (2000).
259. Wilson, R. J. Progress with parasite plastids. *J Mol Biol* **319**, 257-74 (2002).
260. Vivier, E., Petitprez, A. & Landau, I. Observations ultrastructurales sur la sporoblastogenese de l'hemogregarine, Hepatozoon domerguei, Coccidie, Adeleidea. *Protistologica* **8**, 315-333 (1972).
261. Viviere, E. & Hennere, E. Ultrastructure des stades vegetatifs de la coccidie Coelotropha durchoni. *Protistologica* **1**, 89-104 (1965).
262. Schrevel, J. Contribution a l'etude des Selenidiidae parasites d'annelides polychetes II. ultrastructure de quelques trophozoites. *Protistologica* **7**, 101-130 (1971).
263. Cheresh, P., Harrison, T., Fujioka, H. & Haldar, K. Targeting the malarial plastid via the parasitophorous vacuole. *J Biol Chem* **277**, 16265-77 (2002).
264. DeRocher, A., Gilbert, B., Feagin, J. E. & Parsons, M. Dissection of brefeldin A-sensitive and -insensitive steps in apicoplast protein targeting. *J Cell Sci* **118**, 565-74 (2005).
265. Gibbs, S. P. The route of entry of cytoplasmically synthesized proteins into chloroplasts of algae possessing chloroplast ER. *J Cell Sci* **35**, 253-66 (1979).
266. Jarvis, P. & Soll, J. Toc, tic, and chloroplast protein import. *Biochim Biophys Acta* **1590**, 177-89 (2002).
267. Sato, S. & Wilson, R. J. The use of DsRED in single- and dual-color fluorescence labeling of mitochondrial and plastid organelles in Plasmodium falciparum. *Mol Biochem Parasitol* **134**, 175-9 (2004).
268. Sulli, C. & Schwartzbach, S. D. A soluble protein is imported into Euglena chloroplasts as a membrane-bound precursor. *Plant Cell* **8**, 43-53 (1996).
269. Schmidt, R. J., Gillham, N. W. & Boynton, J. E. Processing of the precursor to a chloroplast ribosomal protein made in the cytosol occurs in two steps, one of which depends on a protein made in the chloroplast. *Mol Cell Biol* **5**, 1093-9 (1985).

270. Nassoury, N., Cappadocia, M. & Morse, D. Plastid ultrastructure defines the protein import pathway in dinoflagellates. *J Cell Sci* **116**, 2867-74 (2003).
271. Yung, S., Unnasch, T. R. & Lang-Unnasch, N. Analysis of apicoplast targeting and transit peptide processing in Toxoplasma gondii by deletional and insertional mutagenesis. *Mol Biochem Parasitol* **118**, 11-21 (2001).
272. Ralph, S. A., Foth, B. J., Hall, N. & McFadden, G. I. Evolutionary pressures on apicoplast transit peptides. *Mol Biol Evol* **21**, 2183-94 (2004).
273. Harb, O. S. et al. Multiple functionally redundant signals mediate targeting to the apicoplast in the apicomplexan parasite Toxoplasma gondii. *Eukaryot Cell* **3**, 663-74 (2004).
274. Rigoutsos, I. & Floratos, A. Combinatorial pattern discovery in biological sequences: The TEIRESIAS algorithm. *Bioinformatics* **14**, 55-67 (1998).
275. Rial, D. V., Arakaki, A. K. & Ceccarelli, E. A. Interaction of the targeting sequence of chloroplast precursors with Hsp70 molecular chaperones. *Eur J Biochem* **267**, 6239-48 (2000).
276. Yung, S. C., Unnasch, T. R. & Lang-Unnasch, N. Cis and trans factors involved in apicoplast targeting in Toxoplasma gondii. *J Parasitol* **89**, 767-76 (2003).
277. Schein, A. I., Kissinger, J. C. & Ungar, L. H. Chloroplast transit peptide prediction: a peek inside the black box. *Nucleic Acids Res* **29**, E82 (2001).
278. Dhanasekaran, S., Chandra, N. R., Chandrasekhar Sagar, B. K., Rangarajan, P. N. & Padmanaban, G. Delta-aminolevulinic acid dehydratase from Plasmodium falciparum: indigenous versus imported. *J Biol Chem* **279**, 6934-42 (2004).
279. Waller, R. F., Reed, M. B., Cowman, A. F. & McFadden, G. I. Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *Embo J* **19**, 1794-802 (2000).
280. McMillan, P. J., Stimmler, L. M., Foth, B. J., McFadden, G. I. & Muller, S. The human malaria parasite Plasmodium falciparum possesses two distinct dihydrolipoamide dehydrogenases. *Mol Microbiol* **55**, 27-38 (2005).
281. Cheung, P. Y. et al. Interaction between pyridoxal kinase and pyridoxal-5-phosphate-dependent enzymes. *J Biochem (Tokyo)* **134**, 731-8 (2003).
282. Wrenger, C., Eschbach, M. L., Muller, I. B., Warnecke, D. & Walter, R. D. Analysis of the vitamin B6 biosynthesis pathway in the human malaria parasite Plasmodium falciparum. *J Biol Chem* (2004).
283. Adisa, A. et al. The signal sequence of exported protein-1 directs the green fluorescent protein to the parasitophorous vacuole of transfected malaria parasites. *J Biol Chem* **278**, 6532-42 (2003).
284. Dacie, J. V. & Lewis, S. M. *Practical Haematology* (Churchill Livingstone, Edinburgh, 1975).
285. Saul, A. & Battistutta, D. Analysis of the sequences flanking the translational start sites of Plasmodium falciparum. *Mol Biochem Parasitol* **42**, 55-62 (1990).
286. Roos, D. S., Donald, R. G., Morrissette, N. S. & Moulton, A. L. Molecular tools for genetic dissection of the protozoan parasite Toxoplasma gondii. *Methods Cell Biol* **45**, 27-63 (1994).
287. Joiner, K. A. & Roos, D. S. Secretory traffic in the eukaryotic parasite Toxoplasma gondii: less is more. *J Cell Biol* **157**, 557-63 (2002).
288. Soldati, D. & Boothroyd, J. C. Transient transfection and expression in the obligate intracellular parasite Toxoplasma gondii. *Science* **260**, 349-52 (1993).
289. Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**, 783-95 (2004).
290. Bender, A., van Dooren, G. G., Ralph, S. A., McFadden, G. I. & Schneider, G. Properties and prediction of mitochondrial transit peptides from Plasmodium falciparum. *Mol Biochem Parasitol* **132**, 59-66 (2003).

291. Rudiger, S., Germeroth, L., Schneider-Mergener, J. & Bukau, B. Substrate specificity of the DnaK chaperone determined by screening cellulose-bound peptide libraries. *Embo J* **16**, 1501-7 (1997).
292. Lambros, C. & Vanderberg, J. P. Synchronization of Plasmodium falciparum erythrocytic stages in culture. *J Parasitol* **65**, 418-20 (1979).
293. Prodromou, C. & Pearl, L. H. Recursive PCR: a novel technique for total gene synthesis. *Protein Eng* **5**, 827-9 (1992).
294. Snell, E. E. in *Biochemistry of Vitamin B6 and PQQ* (eds. Marino, G., Sannia, G. & Bossa, F.) (Birkhauser Verlag, Basel, 1994).
295. Gunsalus, I. C., Bellamy, W. D. & Umbreit, W. W. A phosphorylated derivative of pyridoxal as the coenzyme of tyrosine decarboxylase. *J Biol Chem* **155**, 685-686 (1944).
296. Percudani, R. & Peracchi, A. A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Rep* **4**, 850-4 (2003).
297. Eliot, A. C. & Kirsch, J. F. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu Rev Biochem* **73**, 383-415 (2004).
298. Metzler, D. E. *Biochemistry: the Chemical Reactions of Living Cells* (Academic Press, San Diego, CA, 2001).
299. Mehta, P. K. & Christen, P. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes. *Adv Enzymol Relat Areas Mol Biol* **74**, 129-84 (2000).
300. Hayashi, H. Pyridoxal enzymes: mechanistic diversity and uniformity. *J Biochem (Tokyo)* **118**, 463-73 (1995).
301. John, R. A. Pyridoxal phosphate-dependent enzymes. *Biochim Biophys Acta* **1248**, 81-96 (1995).
302. Spenser, I. D. & Hill, R. E. The biosynthesis of pyridoxine. *Nat Prod Rep* **12**, 555-65 (1995).
303. Ehrenshaft, M., Bilski, P., Li, M. Y., Chignell, C. F. & Daub, M. E. A highly conserved sequence is a novel gene involved in de novo vitamin B6 biosynthesis. *Proc Natl Acad Sci U S A* **96**, 9374-8 (1999).
304. Ehrenshaft, M. & Daub, M. E. Isolation of PDX2, a second novel gene in the pyridoxine biosynthesis pathway of eukaryotes, archaebacteria, and a subset of eubacteria. *J Bacteriol* **183**, 3383-90 (2001).
305. Metzler, D. E., Ikawa, M. & Snell, E. E. A general mechanism for vitamin B6-catalysed reactions. *J Am Chem Soc* **76**, 648-652 (1954).
306. Dunathan, H. C. Conformation and reaction specificity in pyridoxal phosphate enzymes. *Proc Natl Acad Sci U S A* **55**, 712-6 (1966).
307. Jansonius, J. N. Structure, evolution and action of vitamin B6-dependent enzymes. *Curr Opin Struct Biol* **8**, 759-69 (1998).
308. Fujii, T. et al. Structure of a NifS homologue: X-ray structure analysis of CsdB, an Escherichia coli counterpart of mammalian selenocysteine lyase. *Biochemistry* **39**, 1263-73 (2000).
309. Kaiser, J. T. et al. Crystal structure of a NifS-like protein from Thermotoga maritima: implications for iron sulphur cluster assembly. *J Mol Biol* **297**, 451-64 (2000).
310. Cupp-Vickery, J. R., Urbina, H. & Vickery, L. E. Crystal structure of IscS, a cysteine desulfurase from Escherichia coli. *J Mol Biol* **330**, 1049-59 (2003).
311. Tirupati, B., Vey, J. L., Drennan, C. L. & Bollinger, J. M., Jr. Kinetic and structural characterization of Slr0077/SufS, the essential cysteine desulfurase from Synechocystis sp. PCC 6803. *Biochemistry* **43**, 12210-9 (2004).

312. Mihara, H. et al. Structure of external aldimine of Escherichia coli CsdB, an IscS/NifS homolog: implications for its specificity toward selenocysteine. *J Biochem (Tokyo)* **131**, 679-85 (2002).
313. Mihara, H., Kurihara, T., Yoshimura, T. & Esaki, N. Kinetic and mutational studies of three NifS homologs from Escherichia coli: mechanistic difference between L-cysteine desulfurase and L-selenocysteine lyase reactions. *J Biochem (Tokyo)* **127**, 559-67 (2000).
314. Gamain, B. et al. Increase in glutathione peroxidase activity in malaria parasite after selenium supplementation. *Free Radic Biol Med* **21**, 559-65 (1996).
315. Fischer, E. H., Kent, A. B., Snyder, E. R. & Krebs, E. G. The reaction of sodium borohydride with muscle phosphorylase. *J Am Chem Soc* **80**, 2906-2907 (1958).
316. Zheng, L., White, R. H., Cash, V. L. & Dean, D. R. Mechanism for the desulfurization of L-cysteine catalyzed by the nifS gene product. *Biochemistry* **33**, 4714-20 (1994).
317. Kessler, D. Slr0077 of Synechocystis has cysteine desulfurase as well as cystine lyase activity. *Biochem Biophys Res Commun* **320**, 571-7 (2004).
318. Dawes, E. A. *Quantitative Problems in Biochemistry* (Longman, London, 1980).
319. Patterson Jnr, G. D. & Pappenhagen, J. M. in *Colormetric Determination of Nonmetals* (eds. Boltz, D. F. & Howell, J. A.) (Wiley-Interscience, New York, N.Y., 1978).
320. Lawrence, N. S., Davis, J. & Compton, R. G. Analytical strategies for the detection of sulfide: a review. *Talanta* **52**, 771-784 (2000).
321. Borup, B. in *Department of Chemistry* (The Pennsylvania State University, 2001).
322. Lacks, S. A. & Springhorn, S. S. Renaturation of enzymes after polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. *J Biol Chem* **255**, 7467-73 (1980).
323. Lawrence, N. S. et al. The Electrochemical Analog of the Methylene Blue Reaction: A Novel Amperometric Approach to the Detection of Hydrogen Sulfide. *Electroanalysis* **12**, 1453-1460 (2000).
324. Kuban, V., Dasgupta, P. K. & Marx, J. N. Nitroprusside and methylene blue methods for silicone membrane differentiated flow injection determination of sulfide in water and wastewater. *Anal Chem* **64**, 36-43 (1992).
325. Lawrence, N. S., Jones, T. G. J., Compton, R. G. & Jiang, L. Recent Advance in the Electrochemical Detection of Sulphide and Sulphhydryl Species. *Chem Res Chinese U* **18**, 161-168 (2002).
326. Siegel, L. M. A Direct Microdetermination for Sulfide. *Anal Biochem* **11**, 126-32 (1965).
327. Miura, N., Yan, Y., Lu, G. & Yamazoe, N. Sensing characteristics and mechanism of hydrogen sulphide sensor using stabilized zirconia and oxide sensing electrode. *Sens Actuators B* **34**, 367-372 (1996).
328. Flint, D. H. Escherichia coli contains a protein that is homologous in function and N-terminal sequence to the protein encoded by the nifS gene of Azotobacter vinelandii and that can participate in the synthesis of the Fe-S cluster of dihydroxy-acid dehydratase. *J Biol Chem* **271**, 16068-74 (1996).
329. Behshad, E., Parkin, S. E. & Bollinger, J. M., Jr. Mechanism of cysteine desulfurase Slr0387 from Synechocystis sp. PCC 6803: kinetic analysis of cleavage of the persulfide intermediate by chemical reductants. *Biochemistry* **43**, 12220-6 (2004).
330. Burri, C. & Brun, R. Eflornithine for the treatment of human African trypanosomiasis. *Parasitol Res* **90 Supp 1**, S49-52 (2003).

331. Greenbaum, D. C. et al. A role for the protease falcipain 1 in host cell invasion by the human malaria parasite. *Science* **298**, 2002-6 (2002).
332. Cleland, W. W. Dithiothreitol, a New Protective Reagent for Sh Groups. *Biochemistry* **35**, 480-2 (1964).
333. Gill, S. C. & von Hippel, P. H. Calculation of protein extinction coefficients from amino acid sequence data. *Anal Biochem* **182**, 319-26 (1989).
334. Burns, J. A., Butler, J. C., Moran, J. & Whitesides, G. M. Selective reduction of disulfides by tris(2-carboxyethyl)phosphine. *J Org Chem* **56**, 2648-2650 (1991).
335. Beinert, H. Semi-micro methods for analysis of labile sulfide and of labile sulfide plus sulfane sulfur in unusually stable iron-sulfur proteins. *Anal Biochem* **131**, 373-8 (1983).
336. Gilboa-Garber, N. Direct spectrophotometric determination of inorganic sulfide in biological materials and in other complex mixtures. *Anal Biochem* **43**, 129-33 (1971).
337. Price, N. C. & Stevens, L. *Fundamentals of Enzymology: the Cell and Molecular Biology of Catalytic Proteins* (Oxford University Press, Oxford, 2003).
338. Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science* **193**, 673-5 (1976).
339. Chen, G. X. & Zolg, J. W. Purification of the bifunctional thymidylate synthase-dihydrofolate reductase complex from the human malaria parasite Plasmodium falciparum. *Mol Pharmacol* **32**, 723-30 (1987).
340. Sorensen, H. P. & Mortensen, K. K. Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli. *Microb Cell Fact* **4**, 1 (2005).
341. Cereghino, G. P., Cereghino, J. L., Ilgen, C. & Cregg, J. M. Production of recombinant proteins in fermenter cultures of the yeast Pichia pastoris. *Curr Opin Biotechnol* **13**, 329-32 (2002).
342. Sirawaraporn, W., Sirawaraporn, R., Cowman, A. F., Yuthavong, Y. & Santi, D. V. Heterologous expression of active thymidylate synthase-dihydrofolate reductase from Plasmodium falciparum. *Biochemistry* **29**, 10779-85 (1990).
343. Cooke, B. M. & Coppel, R. L. Blue skies or stormy weather: what lies ahead for malaria research? *Trends Parasitol* **20**, 611-4 (2004).
344. Baca, A. M. & Hol, W. G. Overcoming codon bias: a method for high-level overexpression of Plasmodium and other AT-rich parasite genes in Escherichia coli. *Int J Parasitol* **30**, 113-8 (2000).
345. Withers-Martinez, C. et al. PCR-based gene synthesis as an efficient approach for expression of the A+T-rich malaria genome. *Protein Eng* **12**, 1113-20 (1999).
346. Prapunwattana, P., Sirawaraporn, W., Yuthavong, Y. & Santi, D. V. Chemical synthesis of the Plasmodium falciparum dihydrofolate reductase-thymidylate synthase gene. *Mol Biochem Parasitol* **83**, 93-106 (1996).
347. Carrio, M. M. & Villaverde, A. Construction and deconstruction of bacterial inclusion bodies. *J Biotechnol* **96**, 3-12 (2002).
348. Carrio, M., Gonzalez-Montalban, N., Vera, A., Villaverde, A. & Ventura, S. Amyloid-like properties of bacterial inclusion bodies. *J Mol Biol* **347**, 1025-37 (2005).
349. Neidhardt, F. C., Ingraham, J. L. & Schaechter, M. *Physiology of the Bacterial Cell: A Molecular Approach* (Sinauer Associates, Sunderland, MA, 1990).
350. Daggett, V. & Fersht, A. The present view of the mechanism of protein folding. *Nat Rev Mol Cell Biol* **4**, 497-502 (2003).

351. Schein, C. H. & Noteborn, M. H. M. Formation of soluble recombinant proteins in Escherichia coli is favored by lower growth temperature. *Biotechnology (N Y)* **6**, 291-4 (1988).
352. Carrio, M. M. & Villaverde, A. Role of molecular chaperones in inclusion body formation. *FEBS Lett* **537**, 215-21 (2003).
353. Mogk, A. et al. Identification of thermolabile Escherichia coli proteins: prevention and reversion of aggregation by DnaK and ClpB. *Embo J* **18**, 6934-49 (1999).
354. Kriek, M., Peters, L., Takahashi, Y. & Roach, P. L. Effect of iron-sulfur cluster assembly proteins on the expression of Escherichia coli lipoic acid synthase. *Protein Expr Purif* **28**, 241-5 (2003).
355. Terpe, K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol* **60**, 523-33 (2003).
356. Smith, J. M. Natural selection and the concept of a protein space. *Nature* **225**, 563-4 (1970).
357. Pedelacq, J. D. et al. Engineering soluble proteins for structural genomics. *Nat Biotechnol* **20**, 927-32 (2002).
358. Maxwell, K. L., Mittermaier, A. K., Forman-Kay, J. D. & Davidson, A. R. A simple in vivo assay for increased protein solubility. *Protein Sci* **8**, 1908-11 (1999).
359. Wilkinson, D. L. & Harrison, R. G. Predicting the solubility of recombinant proteins in Escherichia coli. *Biotechnology (N Y)* **9**, 443-8 (1991).
360. Marston, F. A. & Hartley, D. L. Solubilization of protein aggregates. *Methods Enzymol* **182**, 264-76 (1990).
361. Levinthal, C. in *Mossbauer Spectroscopy in Biological Systems: proceedings of a meeting held at Allerton House, Monticello, Illinois* (eds. DeBrunner, J. T. P. & Munck, E.) (University of Illinois Press, Urbana, Il, 1969).
362. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223-30 (1973).
363. Middelberg, A. P. Preparative protein refolding. *Trends Biotechnol* **20**, 437-43 (2002).
364. Ho, J. G. & Middelberg, A. P. Estimating the potential refolding yield of recombinant proteins expressed as inclusion bodies. *Biotechnol Bioeng* **87**, 584-92 (2004).
365. Das, A. How the phage lambda N gene product suppresses transcription termination: communication of RNA polymerase with regulatory proteins mediated by signals in nascent RNA. *J Bacteriol* **174**, 6711-6 (1992).
366. Davis, G. D., Elisee, C., Newham, D. M. & Harrison, R. G. New fusion protein systems designed to give soluble expression in Escherichia coli. *Biotechnol Bioeng* **65**, 382-8 (1999).
367. Kumar, S. et al. Plant codon optimized cry genes of Bacillus thuringiensis can be expressed as soluble proteins in Escherichia coli BL21 Codon Plus strain as NusA-Cry protein fusions. *J Invertebr Pathol* **88**, 83-6 (2005).
368. Kim, J. S. & Raines, R. T. Ribonuclease S-peptide as a carrier in fusion proteins. *Protein Sci* **2**, 348-56 (1993).
369. Forloni, G. et al. Neurotoxicity of a prion protein fragment. *Nature* **362**, 543-6 (1993).
370. Yamada, R. H., Tsuji, T. & Nose, Y. Uptake and utilization of vitamin B6 and its phosphate esters by Escherichia coli. *J Nutr Sci Vitaminol (Tokyo)* **23**, 7-17 (1977).

371. Belitsky, B. R. Bacillus subtilis GabR, a protein with DNA-binding and aminotransferase domains, is a PLP-dependent transcriptional regulator. *J Mol Biol* **340**, 655-64 (2004).
372. Lo Bello, M. et al. Chemical modification of human placental glutathione transferase by pyridoxal 5'-phosphate. *Biochim Biophys Acta* **1121**, 167-72 (1992).
373. Peterson, E. A. & Sober, H. A. Preparation of crystalline phosphorylated derivatives of vitamin B6. *J Am Chem Soc* **76**, 169-75 (1954).
374. Christensen, H. N. Three Schiff base types formed by amino acids, peptides and proteins with pyridoxal and pyridoxal-5-phosphate. *J Am Chem Soc* **80**, 99-105 (1958).
375. Dayagi, S. & Degani, Y. in *The Chemistry of the Carbon-Nitrogen Double Bond* (ed. Patai, S.) (Interscience Publishers, London, 1970).
376. Finseth, F. & Sizer, I. W. Complexes of pyridoxal phosphate with amino acids, peptides, polylysine, and apotransaminase. *Biochem Biophys Res Commun* **26**, 625-30 (1967).
377. Reyes, A. M., Iriarte, A. & Martinez-Carrion, M. Refolding of the precursor and mature forms of mitochondrial aspartate aminotransferase after guanidine hydrochloride denaturation. *J Biol Chem* **268**, 22281-91 (1993).
378. Herold, M. & Kirschner, K. Reversible dissociation and unfolding of aspartate aminotransferase from Escherichia coli: characterization of a monomeric intermediate. *Biochemistry* **29**, 1907-13 (1990).
379. Relimpio, A., Iriarte, A., Chlebowski, J. F. & Martinez-Carrion, M. Differential scanning calorimetry of cytoplasmic aspartate transaminase. *J Biol Chem* **256**, 4478-88 (1981).
380. Smith, D. B. & Johnson, K. S. Single-step purification of polypeptides expressed in Escherichia coli as fusions with glutathione S-transferase. *Gene* **67**, 31-40 (1988).
381. Walker, P. A. et al. Efficient and rapid affinity purification of proteins using recombinant fusion proteases. *Biotechnology (N Y)* **12**, 601-5 (1994).
382. Frankel, S., Sohn, R. & Leinwand, L. The use of sarkosyl in generating soluble protein after bacterial expression. *Proc Natl Acad Sci U S A* **88**, 1192-6 (1991).
383. Frangioni, J. V. & Neel, B. G. Solubilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. *Anal Biochem* **210**, 179-87 (1993).
384. Filip, C., Fletcher, G., Wulff, J. L. & Earhart, C. F. Solubilization of the cytoplasmic membrane of Escherichia coli by the ionic detergent sodium-lauryl sarcosinate. *J Bacteriol* **115**, 717-22 (1973).
385. Thain, A., Gaston, K., Jenkins, O. & Clarke, A. R. A method for the separation of GST fusion proteins from co-purifying GroEL. *Trends Genet* **12**, 209-10 (1996).
386. Gowda, D. C. & Davidson, E. A. Protein glycosylation in the malaria parasite. *Parasitol Today* **15**, 147-52 (1999).
387. di Guan, C., Li, P., Riggs, P. D. & Inouye, H. Vectors that facilitate the expression and purification of foreign peptides in Escherichia coli by fusion to maltose-binding protein. *Gene* **67**, 21-30 (1988).
388. Pryor, K. D. & Leiting, B. High-level expression of soluble protein in Escherichia coli using a His6-tag and maltose-binding-protein double-affinity fusion system. *Protein Expr Purif* **10**, 309-19 (1997).
389. Ollagnier-De-Choudens, S., Mulliez, E., Hewitson, K. S. & Fontecave, M. Biotin synthase is a pyridoxal phosphate-dependent cysteine desulfurase. *Biochemistry* **41**, 9145-52 (2002).

390. Ollagnier-de-Choudens, S., Mulliez, E. & Fontecave, M. The PLP-dependent biotin synthase from Escherichia coli: mechanistic studies. *FEBS Lett* **532**, 465-8 (2002).
391. Cosper, M. M. et al. Characterization of the cofactor composition of Escherichia coli biotin synthase. *Biochemistry* **43**, 2007-21 (2004).
392. Tse Sum Bui, B., Lotierzo, M., Escalettes, F., Florentin, D. & Marquet, A. Further investigation on the turnover of Escherichia coli biotin synthase with dethiobiotin and 9-mercaptodethiobiotin as substrates. *Biochemistry* **43**, 16432-41 (2004).
393. Jarrett, J. T. The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. *Arch Biochem Biophys* **433**, 312-21 (2005).
394. Alexandrov, A., Dutta, K. & Pascal, S. M. MBP fusion protein with a viral protease cleavage site: one-step cleavage/purification of insoluble proteins. *Biotechniques* **30**, 1194-8 (2001).
395. Routzahn, K. M. & Waugh, D. S. Differential effects of supplementary affinity tags on the solubility of MBP fusion proteins. *J Struct Funct Genomics* **2**, 83-92 (2002).
396. Guruprasad, K., Reddy, B. V. & Pandit, M. W. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Eng* **4**, 155-61 (1990).
397. Tobias, J. W., Shrader, T. E., Rocap, G. & Varshavsky, A. The N-end rule in bacteria. *Science* **254**, 1374-7 (1991).
398. Cserzo, M., Eisenhaber, F., Eisenhaber, B. & Simon, I. On filtering false positive transmembrane protein predictions. *Protein Eng* **15**, 745-52 (2002).
399. Gupta, R. et al. Scanning the available Dictyostelium discoideum proteome for O-linked GlcNAc glycosylation sites using neural networks. *Glycobiology* **9**, 1009-22 (1999).
400. Horton, R. M., Hunt, H. D., Ho, S. N., Pullen, J. K. & Pease, L. R. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-8 (1989).
401. Cordingley, M. G., Callahan, P. L., Sardana, V. V., Garsky, V. M. & Colonno, R. J. Substrate requirements of human rhinovirus 3C protease for peptide cleavage in vitro. *J Biol Chem* **265**, 9062-5 (1990).
402. Zhao, G., Meier, T. I., Hoskins, J. & Jaskunas, S. R. Penicillin-binding protein 2a of Streptococcus pneumoniae: expression in Escherichia coli and purification and refolding of inclusion bodies into a soluble and enzymatically active enzyme. *Protein Expr Purif* **16**, 331-9 (1999).
403. Lin, M. F., Williams, C., Murray, M. V., Conn, G. & Ropp, P. A. Ion chromatographic quantification of cyanate in urea solutions: estimation of the efficiency of cyanate scavengers for use in recombinant protein manufacturing. *J Chromatogr B Analyt Technol Biomed Life Sci* **803**, 353-62 (2004).
404. DiMarchi, R. D. (Eli Lillly and Company, USA, 1986).